



Structure et activité des Archaea planctoniques dans les écosystèmes aquatiques

Mylène Hugoni

► To cite this version:

Mylène Hugoni. Structure et activité des Archaea planctoniques dans les écosystèmes aquatiques. Sciences agricoles. Université Blaise Pascal - Clermont-Ferrand II, 2013. Français. NNT : 2013CLF22388 . tel-01136205

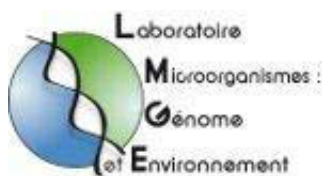
HAL Id: tel-01136205

<https://theses.hal.science/tel-01136205>

Submitted on 26 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE BLAISE PASCAL

Année 2013

Numéro d'ordre : D.U. 2388

**ECOLE DOCTORALE DES SCIENCES DE LA VIE,
SANTÉ, AGRONOMIE ENVIRONNEMENT**
(Numéro 618)

Thèse

Présentée à l'Université Blaise Pascal pour l'obtention du grade de

DOCTEUR D'UNIVERSITE
(Spécialité : **Ecologie Microbienne**)

Soutenue le 31 octobre 2013

Mylène HUGONI

**Structure et activité des *Archaea* planctoniques
dans les écosystèmes aquatiques**

Président :

Didier DEBROAS

Pr, Université Clermont Ferrand II

Rapporteurs :

Céline BROCHIER

Pr, Université Claude Bernard Lyon I

Purificación LOPEZ-GARCIA

DR, Université Paris Sud XI

Membres :

Pierre GALAND

CR, Université Pierre et Marie Curie Paris VI

Laurent TOFFIN

Dr, Ifremer Brest

Isabelle MARY

MCU, Université Clermont Ferrand II

Laboratoire Microorganismes Génome et Environnement – UMR CNRS 6023

Remerciements

Trois années de thèse... comment dire... : peut être : « Ca laisse des marques !!! » et c'est un travail que l'on ne peut pas mener tout seul...

A cet effet, je tiens à remercier dans un premier temps Purificación Lopez-Garcia et Céline Brochier qui ont accepté de rapporter mes travaux, mais aussi Laurent Toffin et Pierre Galand qui ont été examinateurs de mes travaux. Mes remerciements vont également aux deux directeurs d'unité qui m'ont accueilli, M. Amblard et M. Sime-Ngando et mon encadrante, Isabelle, qui a passé son HDR pour moi (enfin en partie !!).

Merci à l'équipe Parasito de ne pas m'avoir gardé en stage de Master 2, ce qui m'a donné l'occasion d'être intégrée à l'équipe MEB, certes au début dans les larmes mais pour mon plus grand bonheur au final !!

Les plus gros des MERCI vont évidemment à mon équipe. Sans le soutien de l'ensemble des membres qui constitue l'équipe MEB je peux affirmer que je n'aurais jamais fini cette thèse. Merci pour tous ces horoscopes, ces mots fléchés, ces blagues à raz les pâquerettes, cette convivialité !! Merci à Didier qui m'a beaucoup appris et épaulé, pas toujours dans le calme mais presque toujours dans la bonne humeur... Je pense aussi évidemment à ma camarade de galère, Najwa, qui tout comme moi n'a qu'une demi-thèse mais un PNAS !!! Un énorme merci d'avoir fait de moi avec Gisèle, ce que l'on peut appeler « une bioinformaticienne en herbe », du moins je crois... Mais sans séquences il n'y a pas de bioinformaticiens, et j'en profite pour remercier très chaleureusement Coco, Anne et les anciens GIIM pour leurs très bons conseils à la pailasse (mais pas que, ils ont aussi joué un rôle de soutien moral) qui sont venus compléter tout ce qu'Agnès m'a appris... Je pense aussi à Piou-piou, Isa Jouan et JC pour leur bonne humeur...

Grosse pensée pour Jonathan, Stéphanie et Marie, avec qui nous avons bravé le froid et les coups de soleil pour aller vider le Pavin et chercher de l'eau toute une année...

Merci à Hélène, Pierre et Zab, collaborateurs et partenaires de Divaqua... mais j'espère bien que ce n'était que le premier de bien d'autres projets que nous mèneront ensemble. Pierre, merci pour ta patience et ton aide (je n'oublierais pas qu'on a passé quasiment un été au téléphone pour combler mes envies de bleu en bordure de papier !!).

Merci aux membres du labo qui m'ont soutenu pendant cette thèse : Fifounette, Cécile, Brigitte, Nathalie, Yvette...

Ensuite, merci à mes amis, évidemment, à propos desquels je n'ai pas besoin d'en dire plus, les gens en savent déjà bien assez !! Emilie, Eric, Benjamin, Anne, Olivier, Elo, Loulou, Nico, Seb, Priscilla, Bruno, Benoît, Stéphanie... Mes amis de Mairie à Leucate : Nelly, Thierry...

Enfin, une dernier merci à ma famille, maman, Pilou, mémés (oui au pluriel !!), papa... et les autres.

Résumé

Les *Archaea* planctoniques contribuent de façon significative aux grands cycles biogéochimiques dans les écosystèmes aquatiques, néanmoins la structure des communautés actives ainsi que leurs variations saisonnières sont encore largement méconnues. En outre, la découverte de l'implication des *Archaea* dans le cycle de l'azote (Ammonia Oxidizing *Archaea* ou AOA), plus particulièrement dans le processus de nitrification a considérablement modifié la perception d'un processus autotrophe réalisé uniquement par des bactéries (Ammonia Oxidizing Bacteria ou AOB). Dans les écosystèmes marins, la large distribution des AOA suggère que ces microorganismes joueraient un rôle prépondérant dans le cycle de l'azote néanmoins, ces observations ne sont pas généralisables à l'ensemble des écosystèmes aquatiques en raison de leur grande diversité et/ou d'un manque d'informations et d'études sur certains d'entre eux.

Ainsi, les objectifs de ce projet étaient i) d'étudier la structure spatiale et temporelle des communautés d'*Archaea* actives dans des écosystèmes aquatiques contrastés en termes d'apports anthropiques et/ou de gradients de salinité (lac, estuaire, milieu côtier); ii) de déterminer la contribution relative des *Archaea* au processus d'oxydation de l'ammonium, en comparaison avec celle des bactéries; et iii) de mieux comprendre les paramètres environnementaux qui pourraient déterminer l'établissement des communautés d'AOA ou d'AOB.

Aquatic *Archaea* are important players among microbial plankton and significantly contribute to biogeochemical cycles, especially nitrogen, but details regarding their community structure and seasonal activity and dynamics remain largely unexplored.

In marine ecosystems, the widespread distribution of Ammonia Oxidizing *Archaea* (AOA) suggests that they probably play a major role in nutrients cycling. However, we cannot generalize these observations to all aquatic ecosystems because of their high diversity and/or a lack of information and studies on these organisms for some of these ecosystems. More precisely, lacustrine and coastal ecosystems were less studied while they are potentially subjected to strong anthropogenic impacts. Moreover, notable differences in terms of diversity and activity between marine and freshwater communities can be expected, considering the specific environmental parameters of each ecosystem.

The objectives of this thesis were: i) to study the archaeal community structure across a temporal scale and assess the diversity of archaeal communities and AOA in diverse aquatic ecosystems along anthropogenic and/or salinity gradient (lacustrine, estuarine and coastal ecosystems); ii) to determine their relative contribution in ammonia oxidation, compared to Ammonia Oxidizing *Bacteria* (AOB) by looking at their spatial and temporal distribution and activity, and iii) to explore more precisely the environmental parameters that could drive AOA and/or AOB establishment.

Liste des Abréviations

16S	16 Sverdberg	MCG	Miscellaneous Crenarchaeotic Group
3HP-4HB	3-hydroxypropionate 4-hydroxybutyrate	MGI	Marine Group I
ADN	Acide Désoxyribonucléique	MGII	Marine Group II
ADNg	ADN Génomique	N₂	Diazote
ADNr	ADN Ribosomique	NGS	Nouvelle Génération de Séquençage
AmoA	Sous-unité A de l'ammonium monooxygénase	NH₂OH	Hydroxylamine
AmoB	Sous-unité B de l'ammonium monooxygénase	NH₄⁺	Ammonium
AmoC	Sous-unité C de l'ammonium monooxygénase	nir	Gène codant la Nitrite Reductase
AOA	Ammonia Oxidizing <i>Archaea</i>	NO	Monoxyde d'azote
AOB	Ammonia Oxidizing <i>Bacteria</i>	NO₂⁻	Nitrites
ARN	Acide Ribonucléique	NO₃⁻	Nitrates
ARNm	ARN messenger	nor	Gène codant la Nitric Oxide Reductase
ARNr	ARN ribosomique	NURM	Complexe N-oxide-Ubiquinone Redox
CNRS	Centre National de la Recherche Scientifique	NXOR	Nitroxyde Oxydoréductase
CO₂	Dioxyde de Carbone	OTU	Operational Taxonomic Unit
CoA	Coenzyme A	PANAM	Phylogenetic Analysis of Next Generation Amplicons
DNRA	Dissimilatory Nitrates Reduction to Ammonium	PCR	Polymerase Chain reaction
EDTA	Ethylene Diamine Tetraacetic Acid	PON	Particulate Organic Nitrogen
GDGT	Glycérol Dialkyl Glycérol Tétratéthers	PPi	Pyrophosphate
HAO	Hydroxylamine Oxidoreductase	QIIME	Quantitative Insights Into Microbial Ecology
HCO₃⁻	Bicarbonates	RC-V	Rice Cluster -V
HNO	Nitroxyde	RDP	Ribosomal Database Project
HNO₂	Acide nitreux	RT-qPCR	Reverse Transcription-PCR
HWCGI	Hot Water Crenarchaeotic Group I	RubisCO	Ribulose 1,5-bisphosphate carboxylase/oxygénase
HWCGIII	Hot Water Crenarchaeotic Group III	SAGMCG-1	South African Gold Mine Group - 1
Km	Constante de Michaelis	SDS	Sodium Dodecyl Sulfate
LDS	Lake Dagow Sediment	ThAOA	AOA Thermophile
MBG-B	Marine Benthic Group B	UreC	Sous-unité A de l'uréase

SOMMAIRE

1	INTRODUCTION GENERALE	1
2	REVUE BIBLIOGRAPHIQUE	7
2.1	Les <i>Archaea</i> dans la classification du vivant	9
2.1.1	Les <i>Euryarchaeota</i>	10
2.1.2	Les <i>Crenarchaeota</i>	12
2.1.3	Les <i>Thaumarchaeota</i>	13
2.1.4	Les <i>Aigarchaeota</i> , les <i>Korarchaeota</i> , les <i>Nanoarchaeota</i>	14
2.2	Distribution des <i>Archaea</i> dans les écosystèmes aquatiques	15
2.2.1	Distribution spatiale des <i>Archaea</i> : répartition cosmopolite ou endémique ?	16
2.2.2	La biosphère rare archéenne : de la théorie à la réalité	17
2.2.3	Distribution des <i>Euryarchaeota</i>	20
2.2.4	Distribution des <i>Thaumarchaeota</i>	20
2.3	Implication des <i>Archaea</i> dans le cycle de l'azote	22
2.3.1	L'azote, un composé essentiel dans les écosystèmes	23
2.3.2	Diversité des <i>Archaea</i> oxydant l'ammonium (AOA)	24
2.4	Métabolismes des <i>Archaea</i> nitrifiantes	27
2.4.1	Éléments intervenant dans la nitrification chez les AOA.....	27
2.4.2	Assimilation autotrophe du carbone	30
2.4.3	Diversité des métabolismes retrouvés chez les AOA.....	32
2.5	Facteurs contrôlant l'abondance et la répartition des AOA en milieu aquatique.....	33
2.5.1	Influence de la concentration en ammonium	33
2.5.2	Importance de la salinité et du pH.....	34
2.6	Conclusions et hypothèses	35
3	SITES D'ETUDE ET APPROCHES METHODOLOGIQUES	37
3.1	Description des sites d'étude	39
3.1.1	Les milieux d'eau douce	39
3.1.1.1	Le lac d'Aydat.....	39
3.1.1.2	Le lac Pavin.....	40
3.1.1.3	Le réservoir de la Sep.....	41
3.1.1.4	Le lac du Bourget.....	42
3.1.2	Estuaire et milieu côtier.....	43
3.1.2	La baie de Banyuls sur Mer.....	43
3.1.3	L'estuaire de la Charente.....	44
3.2	Approches d'écologie moléculaire.....	45
3.2.1	Conservation des échantillons et extraction des acides nucléiques.....	45

3.2.2	Amplification par PCR des gènes cibles	46
3.2.2.1	La PCR quantitative	48
3.2.2.2	Le pyroséquençage d'amplicons ou métagénétique.....	49
3.2.3	Traitement bioinformatique des données de NGS	51
4	CHAPITRE 1 : Structure des communautés d'<i>Archaea</i> actives dans les écosystèmes aquatiques	53
4.1	Contexte et objectifs.....	55
4.2	Article 1 : Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters.....	56
4.3	Article 2 : New insights into active archaeal groups in oxygen depleted zones of lacustrine ecosystems	73
5	CHAPITRE 2 : Impact de gradients environnementaux sur la diversité spécifique et fonctionnelle des <i>Archaea</i>	103
5.1	Contexte et objectifs.....	105
5.2	Article 3 : Dynamics of ammonia-oxidizing <i>Archaea</i> and <i>Bacteria</i> in contrasted freshwater ecosystems.....	106
5.3	Article 4 : Temporal and spatial dynamics of active prokaryotes nitrifiers and archaeal communities from river to sea	121
6	DISCUSSION ET PERSPECTIVES	151
6.1	Apports et biais des études de métagénétique	154
6.1.1	Spécificité des amorces	154
6.1.2	Nettoyage des séquences et clusterisation.....	156
6.1.3	Reconstruction phylogénétique	156
6.2	Composition et activité de la communauté archéenne aquatique.....	157
6.2.1	Ecotypes liés aux groupes MGI et MGII	159
6.2.2	Des groupes très actifs au métabolisme peu connu.....	166
6.3	Importance de la biosphère rare archéenne dans les écosystèmes aquatiques	167
7	REFERENCES BIBLIOGRAPHIQUES.....	177
8	<i>CURRICULUM VITAE</i>	191

INTRODUCTION GENERALE

L'étude et la compréhension du fonctionnement des écosystèmes aquatiques, au regard notamment des flux d'énergie et de matière qui les caractérisent, ont conduit à théoriser et modéliser le fonctionnement de ces systèmes. Si les réseaux trophiques étaient représentés, dans un premier temps, par une chaîne linéaire allant de l'assimilation photosynthétique par le phytoplancton vers des transferts de matière et d'énergie aux maillons trophiques supérieurs (zooplancton et poissons), les avancées scientifiques ont mis en lumière l'importance du compartiment microbien, *via* le concept de boucle microbienne (Azam *et al.*, 1983). Ainsi, la caractérisation des interactions trophiques et les flux de matière dépendent de la connaissance des communautés planctoniques qui en sont à l'origine, parmi lesquelles les *Archaea*.

La présence d'*Archaea* dans des environnements extrêmes, tels que des sources chaudes (Woese and Fox, 1977) ou des écosystèmes digestifs (Paynter and Hungate, 1968) est avérée depuis de nombreuses années. Cependant, leur présence dans les écosystèmes aquatiques oxygénés a été mise en évidence en 1992 seulement (DeLong, 1992; Fuhrman *et al.*, 1992). L'utilisation croissante des techniques moléculaires a permis de suggérer qu'elles présentaient une large distribution dans la plupart des environnements (DeLong and Karl, 2005), pourtant certains écosystèmes tels que les milieux lacustres, ont été peu étudiés comparativement aux milieux marins, et la biogéographie de ces microorganismes reste à étudier.

D'autre part, les données disponibles sur la diversité et le rôle de ces *Archaea* aquatiques découlent des connaissances scientifiques des groupes les plus abondants dans les écosystèmes. L'existence de microorganismes rares est pourtant reconnue depuis la mise en œuvre des techniques culturelles mais le nombre important d'OTUs qui les représente a longtemps été négligé comme en témoigne le nombre d'études croissant tendant à montrer une grande richesse d'espèces rares (Campbell *et al.*, 2011; Pedros-Alio, 2006; Pedros-Alio, 2012; Sogin *et al.*, 2006). Néanmoins, la majeure partie de ces études ont été conduites sur la fraction bactérienne et peu se sont focalisées sur la structuration et l'activité de la biosphère rare archéenne (Galand *et al.*, 2009).

Les études conduites en milieu aquatique oxygéné ont montré que les *Archaea* jouaient un rôle majeur dans les cycles biogéochimiques *via* leur implication dans les cycles du carbone et de l'azote. Ces microorganismes pourraient donc avoir un rôle non négligeable dans le fonctionnement de ces écosystèmes en prenant part à des processus historiquement décrits chez les protéobactéries dans le cas de l'oxydation de l'ammonium. Ce processus qui est également réalisé par les *Archaea*, a lieu aussi bien dans les milieux marins que dans les milieux d'eau douce (Auguet *et al.*, 2011; Lliros *et al.*, 2008) et implique des gènes codant

des ammonium monooxygénases (gènes *amoA*) spécifiques de ces microorganismes (Herrmann *et al.*, 2008). Néanmoins de nombreuses questions demeurent encore aujourd'hui en suspens, notamment sur la contribution relative des Ammonia Oxidizing *Archaea* (AOA) par rapport à celle des Ammonia Oxidizing *Bacteria* (AOB), laissant ouvert un vaste champ d'investigation. En effet, l'abondance des gènes *amoA* d'*Archaea* est quelquefois supérieure à celle des gènes *amoA* d'AOB dans certains milieux marins et terrestres (Leininger *et al.*, 2006; Wuchter *et al.*, 2006), mais cette tendance ne peut être généralisée à l'ensemble des écosystèmes en raison de leur grande diversité et/ou du manque de données les concernant. Par ailleurs, peu de travaux ont porté sur l'étude conjointe de l'abondance et de l'activité des AOA et des AOB au sein d'un même écosystème. Généralement, seul un des deux groupes était considéré, ce qui ne permettait pas d'identifier les facteurs susceptibles d'impacter à la fois la diversité et l'activité de chacun de ces acteurs dans le processus de nitrification. L'étude des communautés nitrifiantes est aujourd'hui nécessaire car elles peuvent être impliquées dans les phénomènes d'eutrophisation *via* le recyclage et la mise à disposition de composés azotés minéraux aux producteurs primaires. Ainsi, un apport trop important d'azote couplé à un apport en phosphore excessif peut provoquer des perturbations du milieu à l'origine de conséquences écologiques, socio-économiques et sanitaires non négligeables.

Afin de bien cerner le contexte de ce travail qui visait à approfondir notre connaissance des communautés archéennes (en termes de composition, de distribution ou d'écologie), et de leur rôle dans le fonctionnement des écosystèmes, ce mémoire, rédigé sous la forme d'une thèse sur articles, s'articule autour de quatre parties :

- La première partie du manuscrit est consacrée à une revue bibliographique qui retrace les données actuelles disponibles sur la structuration et la diversité des communautés d'*Archaea* en milieu aquatique, ainsi que sur les communautés impliquées dans le cycle biogéochimique de l'azote.

- Puis, les différentes techniques d'écologie moléculaire employées et les sites d'études utilisés lors de ce travail seront plus largement détaillés. En effet, nous avons travaillé sur des milieux naturels très différents : lacustre, côtier et estuarien, qui nous ont permis de bénéficier pour certains de collaborations avec des Observatoires Microbiologiques impliqués dans des observations à long terme. En outre, ce choix nous a permis de travailler dans des milieux

présentant des gradients anthropiques ou encore des gradients de salinité, facteurs susceptibles d'impacter les communautés d'*Archaea* et leur activité.

- Le premier chapitre présente les principaux résultats obtenus au cours de deux études. Ces travaux nous ont permis de caractériser la structure des communautés actives archéennes au cours d'une étude à long terme menée en milieu côtier. Nous avons également pu évaluer finement l'importance et l'activité de la biosphère rare des *Archaea*. Dans le second volet de ce chapitre, une étude conduite dans des milieux peu étudiés, les milieux lacustres, sera présentée.

- Le second chapitre s'articule autour de deux articles portant sur l'évaluation de l'impact des facteurs environnementaux sur l'activité des communautés d'*Archaea* et, spécifiquement sur l'activité des AOA. Ainsi, la première étude a été conduite sur des lacs de statuts trophiques différents tandis que la seconde étude a été réalisée le long d'un estuaire, écosystème modèle pour les transitions chimiques et microbiennes.

- Enfin, la dernière partie de ce manuscrit sera consacrée à une discussion générale articulée autour de trois parties : les apports et les biais des études de métagénétique, la composition et l'activité des communautés d'*Archaea* aquatiques et enfin l'importance de la biosphère rare archéenne dans les milieux aquatiques.

REVUE BIBLIOGRAPHIQUE

2.1 Les *Archaea* dans la classification du vivant

Pendant longtemps, les microorganismes ont été classés en fonction de critères phénotypiques tels que leur morphologie, leur physiologie ou encore leur pathogénicité. Alors que ces critères permettaient une classification, ils n'autorisaient cependant pas la prise en compte des relations évolutives existant entre les différents groupes taxonomiques. En effet, il s'avère que les caractères phénotypiques sont peu résolutifs pour les niveaux taxonomiques les plus fins. Ceci a été confirmé par le travail pionnier de Woese et Fox en 1977, qui ont élaboré pour la première fois une classification du vivant sur la base d'un critère phylogénétique, à savoir la séquence des gènes codant les ARN ribosomiques (Woese and Fox, 1977), une molécule présente chez tous les organismes vivants. Les résultats qui ont découlé de ces travaux ont bouleversé la classification initiale. Ce travail a fait l'objet de nombreuses polémiques (Mayr, 1998) et la distribution des organismes vivants en trois domaines distincts (i.e. *Eucarya*, *Bacteria* et *Archaea* (Figure 1)), n'a été acceptée que tardivement par la communauté scientifique (Woese, 1990, 1994).

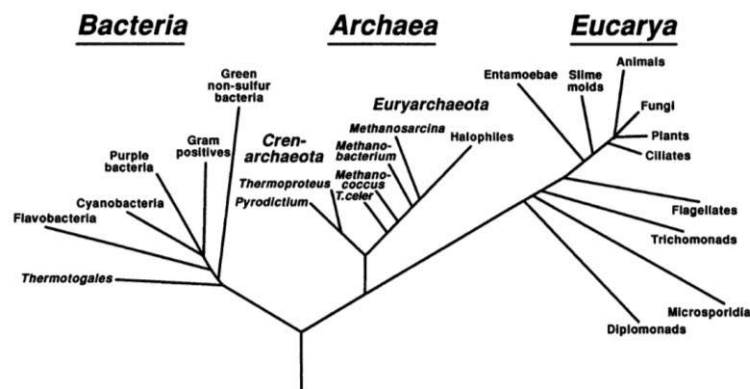


Figure 1 : Arbre phylogénétique construit à partir des séquences codant la petite sous-unité de l'ARN ribosomique (ARNr 16S pour les procaryotes et 18S pour les eucaryotes) et traduisant l'existence de trois domaines du vivant, d'après Woese (1990 et 1994).

Dans l'arbre du vivant déterminé sur la base du gène codant les ARNr, les *Archaea* se détachent et forment un embranchement à part entière. Celui-ci rassemble des organismes unicellulaires porteurs à la fois de caractéristiques bactériennes et eucaryotes (Pace, 2006). Les *Archaea* sont similaires aux bactéries pour de nombreux traits, essentiellement en relation avec la taille cellulaire et génomique, le coefficient de sédimentation des ADNr (16S et 23S),

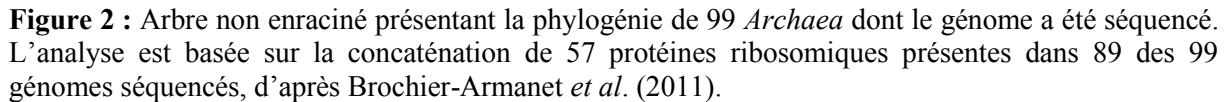
le mode de division cellulaire et l'absence de noyau. D'autres aspects les rapprochent cependant davantage des eucaryotes. Il s'agit principalement des mécanismes et des protéines impliqués dans les processus liés à l'ADN tels que la réplication, la transcription, la traduction et les mécanismes de réparation (Forterre *et al.*, 2002).

Quelques particularités sont, quant à elles, propres aux *Archaea* et permettent de les différencier des deux autres domaines. Pour la majorité, ces particularités portent sur la structure et la composition de leur membrane et de leur paroi cellulaire (Brown and Doolittle, 1997). Ainsi, leur paroi est dépourvue de peptidoglycane, composant retrouvé habituellement dans celle de bactéries. Quant à leur membrane, les phospholipides la constituant sont faits de lipides comportant de longues chaînes latérales isoprénoïdes, possédant une stéréochimie propre (Matsumi *et al.*, 2010). En effet, chez les *Archaea*, l'agencement des chaînes latérales C₂₀ isoprénoïdes conduit à la formation d'une monocouche lipidique, contrairement à la bicouche retrouvée chez les bactéries. Enfin, le cœur apolaire de la membrane cellulaire est dominé par la présence de glycérol dialkyl glycérol tétraéthers (GDGT).

Les premières analyses phylogénétiques basées sur les séquences d'ADNr 16S ont permis de définir deux *phyla* archéens, les *Euryarchaeota* et les *Crenarchaeota* (Woese *et al.*, 1990). Ces données, couplées à l'analyse récente de génomes, ont permis de proposer quatre nouveaux *phyla* au sein de ce domaine : les *Thaumarchaeota*, les *Korarchaeota*, les *Nanoarchaeota* et les *Aigarchaeota* (Figure 2).

2.1.1 Les *Euryarchaeota*

Le nom *Euryarchaeota* a pour origine le mot grec « Eurys » qui signifie large. Ceci est dû au fait que ce *phylum* regroupe des *Archaea* physiologiquement diverses parmi lesquelles beaucoup sont endémiques des milieux extrêmes. Il comprend à la fois des espèces hyperthermophiles, thermophiles et mésophiles. Ce *phylum* inclut des organismes méthanogènes (*Methanobacteriales*, *Methanococcales*...), des halophiles extrêmes appartenant aux *Halobacteriales*, des hyperthermophiles (*Thermococcales*, *Archaeoglobales*) ou encore des acidophiles (*Thermoplasmatales*).



De plus, les *Euryarchaeota* sont capables de se développer dans des milieux très contrastés comme, par exemple la zone pélagique de milieux marins colonisée par les *Euryarchaeota* des Marine Group II et III ou encore les milieux d'eau douce où des *Euryarchaeota* des clusters LDS (Lake Dagow Sediment, (Glissman *et al.*, 2004)) ou RC-V (Rice Cluster-V, (Galand *et al.*, 2006; Herfort *et al.*, 2009)) ont pu être identifiés.

Dans ce *phylum*, certains groupes d'*Archaea* occupent des fonctions clés bien identifiées comme les méthanogènes, tandis que d'autres, comme les *Euryarchaeota* du Marine Group II, des clusters LDS ou RC-V n'ont aucun métabolisme clairement défini alors qu'ils peuvent représenter une part non négligeable de la fraction procaryotique aquatique (Barberan *et al.*, 2011; Karner *et al.*, 2001).

2.1.2 Les *Crenarchaeota*

Les *Crenarchaeota* sont physiologiquement moins diversifiées que les *Euryarchaeota*. Ce *phylum* a été historiquement défini à partir d'espèces thermophiles et hyperthermophiles d'où le nom « Crene » pour source ou fontaine en grec, en référence aux sources chaudes d'où elles ont été initialement isolées. Certaines d'entre elles présentent les plus hautes températures de croissance jamais observées (113°C pour *Pyrolobus fumarii* (Blochl *et al.*, 1997)). La plupart de ces hyperthermophiles sont des autotrophes chimiolithotrophes. Ces dernières appartiennent principalement à deux ordres, les *Desulfurococcales* et les *Thermoproteales* (Figure 2).

Des études menées sur des environnements aquatiques et basées sur les techniques moléculaires ont permis de caractériser de nouvelles séquences de gènes d'ADNr 16S d'*Archaea* (DeLong, 1992; Fuhrman *et al.*, 1992). Une partie de ces séquences a dans un premier temps été affiliée aux *Crenarchaeota*, avant d'être reclassée dans un nouveau *phylum*, celui des *Thaumarchaeota*, sur la base de données génomiques (Brochier-Armanet *et al.*, 2008). D'autres séquences, en revanche, ont été conservées au sein des *Crenarchaeota* pour former les groupes Miscellaneous Crenarchaeotic Group (MCG) ou encore Marine Benthic Group B (MBG-B). Cependant, le rôle de ces *Crenarchaeota* dans le fonctionnement des écosystèmes reste encore énigmatique.

2.1.3 Les *Thaumarchaeota*

Comme décrit précédemment, le *phylum* des *Thaumarchaeota* a d'abord été défini à partir de séquences environnementales d'ADNr 16S. Par la suite, quelques espèces ont été cultivées, pour la plupart en condition non axénique (Spang *et al.*, 2012). Grâce aux études de génomique comparative et phylogénétiques menées sur les gènes d'ADNr 16S et 23S et sur des protéines ribosomiques (Brochier-Armanet *et al.*, 2008; Spang *et al.*, 2010), l'existence de ce nouveau *phylum* a été validée. Sur la base des gènes d'ADNr 16S, plusieurs groupes de *Thaumarchaeota* ont été définis (Figure 3). Parmi ces groupes pour lesquels il existe un représentant cultivé figurent :

- le Marine Group I (MGI) également appelé groupe 1.1a, représenté par *Cenarchaeum symbiosum* (Hallam *et al.*, 2006a; Preston *et al.*, 1996), *Nitrosopumilus maritimus* (Konneke *et al.*, 2005; Martens-Habbenha *et al.*, 2009), et *Candidatus Nitrosoarchaeum limnia* (Blainey *et al.*, 2011; Mosier *et al.*, 2012) ;
- un groupe-sœur du groupe 1.1a, appelé South African Gold Mine Crenarchaeotic Group 1 (SAGMCG-1), représenté par *Nitrosotalea devanatterra* (Lehtovirta-Morley *et al.*, 2011) ;
- le groupe 1.1b représenté par *Nitrososphaera gargensis* (Hatzenpichler *et al.*, 2008) et *Nitrososphaera viennensis* (Tourna *et al.*, 2011) ;
- le groupe Hot Water Crenarchaeotic Group III (HWCGIII) représenté par *Nitrosocaldus yellowstonii* (de la Torre *et al.*, 2008), lignée thermophile de *Thaumarchaeota* (ThAOA).

Alors que les séquences affiliées aux *Thaumarchaeota* n'étaient jusqu'à maintenant retrouvées qu'en milieux mésophiles, leur caractérisation récente dans des milieux thermophiles (groupe HWCGIII (de la Torre *et al.*, 2008)) suggère que l'écologie de ce *phylum* est beaucoup plus complexe que présumée. Par ailleurs, l'intérêt de la communauté scientifique pour ce *phylum* s'est largement accru depuis la découverte de leur capacité à oxyder l'ammonium dans les écosystèmes aquatiques, terrestres ou encore dans les sources hydrothermales, démontrant ainsi pour la première fois le potentiel nitrifiant des *Archaea*. Une recherche plus approfondie sur les *Thaumarchaeota*, prenant en compte les sous-groupes non cultivés, s'avère donc essentielle pour mieux caractériser la biologie, l'écologie et l'évolution des espèces composant ce *phylum*.

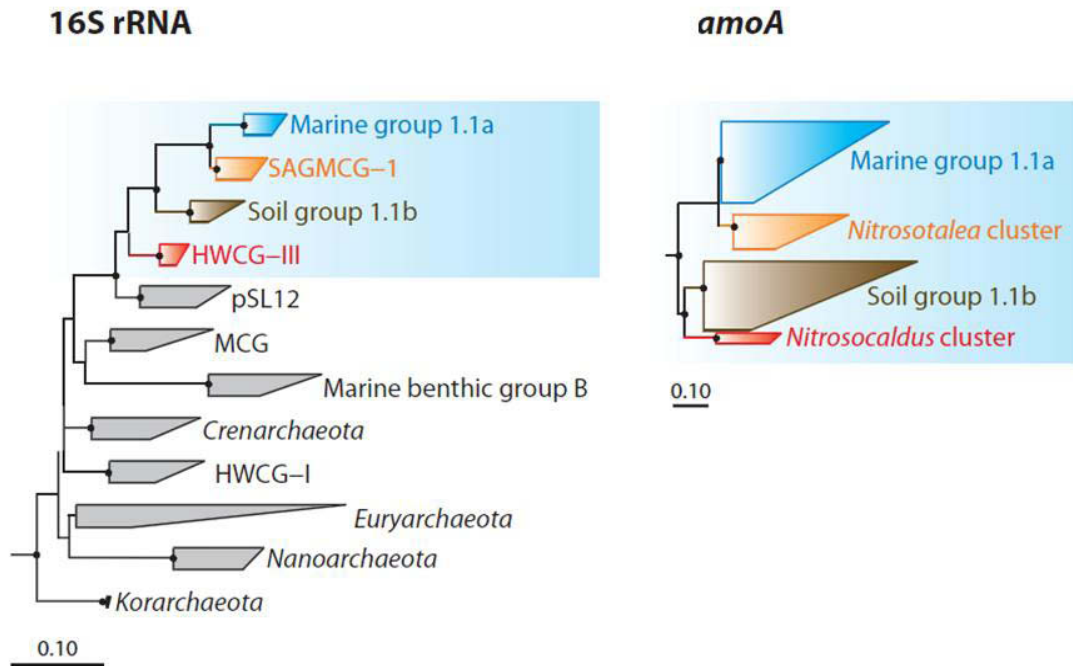


Figure 3 : Phylogénies des *Thaumarchaeota* basées sur l'analyse des séquences d'ADNr 16S (594 positions nucléotidiques) et des gènes *amoA* archéens identifiés pour ce groupe (1265 positions nucléotidiques), d'après Stahl et de la Torre (2012). Pour les principaux groupes, une affiliation taxonomique similaire est observée quel que soit le marqueur utilisé.

2.1.4 Les *Aigarchaeota*, les *Korarchaeota*, les *Nanoarchaeota*

Le domaine des *Archaea* se compose également de trois *phyla* pour lesquels peu de données sont encore aujourd'hui disponibles.

Le *phylum* des *Aigarchaeota* a été proposé sur la base de l'analyse d'un génome composite, celui de *Candidatus Caldiarchaeum subterraneum*, reconstruit à partir de données de métagénomique. Cette espèce isolée d'un tapis microbien, se développe dans les sources géothermales profondes (Nunoura *et al.*, 2011). Les auteurs suggèrent que ce microorganisme pourrait être chimiolithotrophe et utiliser le dihydrogène ou le monoxyde de carbone pour réduire l'oxygène, le nitrate ou le nitrite. Les analyses phylogénétiques montrent que *Candidatus Caldiarchaeum subterraneum* est un membre du Hot Water Crenarchaeotic Group I (HWCGI), groupe également composé d'organismes potentiellement thermophiles détectés dans des environnements tels que les sources chaudes ou les écosystèmes profonds hydrothermaux (de la Torre *et al.*, 2008). Malgré ces données, la position du *phylum* des *Aigarchaeota* (Figure 2), proche des *Thaumarchaeota* et des *Crenarchaeota* dans l'arbre phylogénétique (Nunoura *et al.*, 2011), est encore aujourd'hui très incertaine et fait l'objet de

vifs débats (Brochier-Armanet *et al.*, 2011; Pester *et al.*, 2011). En effet, la distance très faible les séparant des *Thaumarchaeota*, pourrait signifier une position relativement basale au sein des *Thaumarchaeota* (Brochier-Armanet *et al.*, 2011).

Le *phylum* des *Korarchaeota* a, quant à lui, été mis en évidence sur la base de séquences de gènes d'ADNr 16S isolées d'une source d'eau chaude du Parc de Yellowstone (Barns *et al.*, 1996). Il émerge près de la racine de l'embranchement des *Archaea*. Depuis ces découvertes, le génome d'une espèce a été séquencé. Il s'agit de celui de *Candidatus Korarchaeum cryptofilum*, souche issue de cultures d'enrichissement provenant d'échantillons issus de sédiments marins et de sources chaudes (Elkins *et al.*, 2008). L'analyse des séquences composant ce génome indique la présence de gènes *Crenarchaeota*-like (notamment des sous-unités de l'ARN polymérase) et *Euryarchaeota*-like (maturation des ARNt, réplication/réparation de l'ADN et division cellulaire), confortant l'idée qu'il s'agit bien d'un *phylum* à part entière.

Enfin, le troisième *phylum*, celui des *Nanoarchaeota*, a été découvert en 2002 et ne compte aujourd'hui qu'un seul genre, celui des *Nanoarchaeum* (Huber *et al.*, 2002). L'unique espèce affiliée à ce genre, *Nanoarchaeum equitans*, est de petite taille (400 nm de diamètre) et possède un génome de taille réduite (0.5 Mb). Elle vit de façon symbiotique avec une espèce de *Crenarchaeota* thermophile, *Ignococcus hospitalis* (Podar *et al.*, 2008). Cette association, mettant en jeu deux organismes unicellulaires, est la plus simple décrite à ce jour.

2.2 Distribution des *Archaea* dans les écosystèmes aquatiques

Les études menées ces dernières années montrent que les *Archaea* sont présentes de façon ubiquitaire dans les écosystèmes aquatiques, leur abondance ayant même été estimée entre 1 et 10% de la population procaryotique dans les lacs (Keough *et al.*, 2003) et à près de 10^{28} cellules dans l'océan mondial (Karner *et al.*, 2001). Les zones oxiques de ces systèmes sont colonisées principalement par des espèces affiliées aux *Euryarchaeota* et aux *Thaumarchaeota*, très peu de *Crenarchaeota* étant détectées. Cependant, de nombreuses questions relatives aux patrons de distribution de ces microorganismes et aux facteurs qui les déterminent sont actuellement au cœur des problématiques écologiques.

Le nombre croissant d'études de génomique environnementale, rendues possibles essentiellement par l'essor des nouvelles techniques de séquençage (NGS), a permis d'apporter de nouvelles connaissances sur la présence d'*Archaea* dans les écosystèmes

aquatiques. Cependant, les études visant à caractériser leur patron de distribution (en termes de distribution spatiale mais aussi d'abondance) restent marginales comparées au nombre d'études menées sur les communautés bactériennes. De plus, elles concernent essentiellement la fraction la plus représentée de la biosphère archéenne. Or, une étude conduite en 2009 sur un milieu marin a démontré que le nombre d'OTUs dites « rares » affiliées aux *Archaea* était important suggérant un rôle jusqu'alors peu connu de cette composante, dans le fonctionnement des écosystèmes aquatiques (Galand *et al.*, 2009).

2.2.1 Distribution spatiale des *Archaea* : répartition cosmopolite ou endémique ?

Durant les dernières décennies, les études d'écologie microbienne portant sur la distribution des microorganismes dans l'environnement ont abouti à l'émergence de plusieurs concepts. Le premier proposé par Lourens Baas-Becking était basé sur l'hypothèse du "Everything is everywhere, the environment selects" (Baas-Becking, 1934). Dans ce concept, l'auteur suggère une répartition cosmopolite de tous les microorganismes à travers le globe avec une sélection des espèces par le milieu et les paramètres environnementaux, aucune barrière géographique ne limitant leur distribution (Finlay, 2002).

A ce concept s'oppose celui selon lequel la capacité de dispersion géographique des microbes ne serait pas illimitée, mais au contraire, elle varierait fortement selon les espèces en fonction de leur aptitude à survivre lors de migrations, à former des structures de résistance (spores ou kystes) ou à se développer dans un nouvel espace (Galand *et al.*, 2009; Martiny *et al.*, 2006). Ainsi, même si certaines espèces sont cosmopolites, d'autres seraient en revanche endémiques à certaines régions du globe. Dans ce cas, les patrons de distribution des microorganismes seraient conditionnés par les facteurs environnementaux et/ou les paramètres géographiques physiques. Ce concept est supporté par plusieurs travaux de recherche parmi lesquels certains menés sur des écosystèmes aquatiques. Certains travaux suggèrent que la biogéographie des microorganismes serait déterminée par les facteurs environnementaux (Logue *et al.*, 2008), cependant, d'autres montrent que la distance séparant les écosystèmes, et donc l'intensité des échanges d'espèces entre eux, aurait une influence sur leur composition en espèces, et cela indépendamment des autres variables environnementales (Yannarell and Triplett, 2004). Ceci a également été suggéré par le biais de travaux conduits sur des sources d'eau chaude qui incriminent davantage l'éloignement

géographique des sources entre elles pour expliquer les différences observées entre les populations d'*Archaea* appartenant au genre *Sulfolobus* (Whitaker *et al.*, 2003).

Malgré un nombre d'études croissant, peu d'entre elles se sont focalisées sur la distribution des *Archaea* en milieu aquatique et, même si les nombreuses études environnementales tendent à prouver que ces microorganismes sont retrouvés de façon ubiquitaire dans tous les écosystèmes, on ne sait que peu de choses concernant la répartition et l'abondance des différents *phyla* et des clusters qui les constituent. Au même titre, les travaux visant à caractériser la structure des populations d'*Archaea* sont basés essentiellement sur l'étude des *taxa* abondants, alors même que l'importance et le rôle potentiel de la biosphère rare est à l'heure actuelle l'une des grandes questions en écologie microbienne.

2.2.2 La biosphère rare archéenne : de la théorie à la réalité

La détermination de la structure des communautés microbiennes *via* la caractérisation de la richesse, de l'abondance ou de la diversité dans un écosystème est un enjeu central en écologie. La présence de microorganismes rares au sein des communautés microbiennes est avérée depuis la mise en œuvre des techniques d'écologie moléculaires classiques, pourtant leur richesse a dû être reconsidérée au cours des dernières années. En effet, dans les écosystèmes aquatiques, les estimations en terme de richesse étaient jusqu'à récemment de moins de 200 OTUs bactériennes pour 90% des banques de clones (Kemp and Aller, 2004). Cependant, avec le développement des approches de séquençage de nouvelle génération (NGS), une fraction rare beaucoup plus vaste a été mise en évidence (Figure 4).

Plusieurs stratégies reposant sur les outils dits « haut débit » ont été développées pour explorer les communautés microbiennes. Des approches de métagénomique couplant les techniques de NGS à l'amplification ciblée de gènes marqueurs tels que celui codant la petite sous-unité ribosomique (ARNr 16S) ont été mises en place et ont montré leur acuité à décrypter la richesse d'un milieu avec beaucoup de précision (Caporaso *et al.*, 2012; Huber *et al.*, 2007; Sogin *et al.*, 2006).

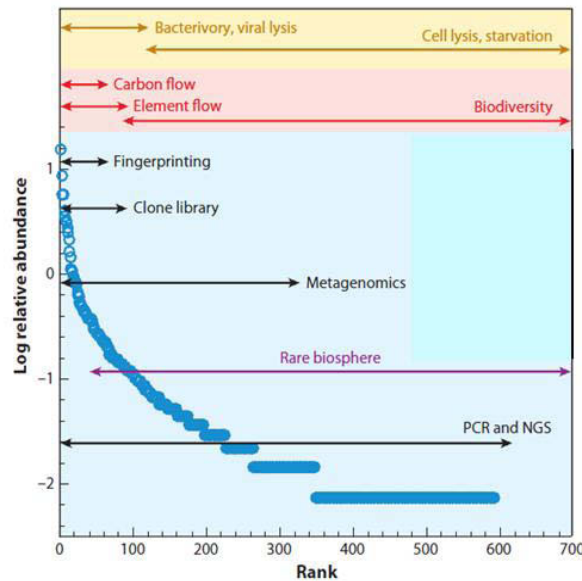


Figure 4 : Schéma représentant la courbe rang-abondance établie pour les bactéries et les différents facteurs les contrôlant. En jaune : les facteurs de perte ; en rouge : les facteurs écologiques et en noir : la profondeur de séquençage permise par les différentes techniques moléculaires. La fraction rare est représentée en violet, d'après Pedros-Alio (2012).

Les résultats issus de cette stratégie montrent un partage des communautés microbiennes en deux fractions distinctes (Figure 5) :

- celle constituée de peu d'espèces mais chacune représentée par un grand nombre d'individus dits *taxa* abondants dont le maintien se ferait par une croissance active. Ces *taxa*, très adaptés à leur environnement, seraient régulés essentiellement par des phénomènes de prédation et de lyse virale (Campbell et al., 2011) ;
- celle, au contraire, constituée d'un grand nombre d'espèces chacune représentée par un faible nombre d'individus, qui formerait la « biosphère rare ». Celle-ci viendrait compléter la diversité présente dans l'écosystème (Sogin et al., 2006).

Néanmoins, la définition de « rareté » reste encore aujourd'hui relativement subjective dans la mesure où le seuil appliqué pour considérer qu'une espèce est rare dans un milieu donné est extrêmement variable d'une étude à l'autre, compris entre <1% (Campbell et al., 2011) et <0.01% (Galand et al., 2009; Mangot et al., 2012). Le seuil choisi doit être basé sur la représentativité du jeu de données mais aussi et surtout, il ne peut être choisi que lorsque l'effort de séquençage nécessaire pour décrire la totalité de la diversité est suffisant, en d'autres termes lorsque la saturation des courbes de raréfaction est atteinte.

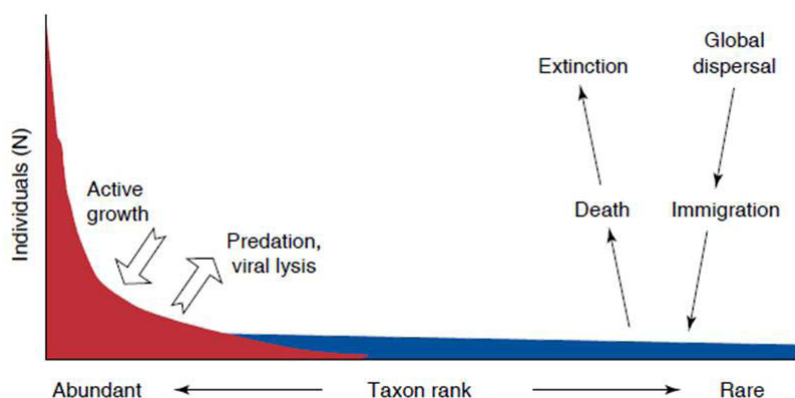


Figure 5 : Représentation schématique d'une courbe rang-abondance représentant le nombre de *taxa* en fonction de leur rang taxonomique et de leur abondance. La courbe représente la biodiversité et postule de l'existence de deux fractions microbiennes, une composée des taxons abondants en rouge et l'autre, des taxons rares, en bleu, d'après Pedros-Alio (2006).

La découverte de cette biosphère rare au sein des communautés microbiennes fait l'objet de débats comme en témoignent certains articles (Huse *et al.*, 2010; Kunin *et al.*, 2009; Quince *et al.*, 2009) et différentes hypothèses. L'une d'entre elle suggère que certains *taxa* rares pourraient passer d'un statut de dormance à actifs et ainsi voir le nombre d'individus les représentant augmenter, en fonction des paramètres environnementaux (Pedros-Alio, 2006; Pedros-Alio, 2012). Cette théorie, dite de « Seed Bank » suggère par conséquent que ces *taxa* rares constitueraient un réservoir de diversité métabolique. En revanche, ce modèle d'alternance rares/dominants s'oppose à une autre vision découlant des travaux de Galand *et al.* (2009a) menés sur la fraction bactérienne et archéenne, qui montrent que certains de ces microorganismes détectés comme rares, restent toujours rares quelles que soient les conditions environnementales et qu'ils sont spécifiques d'une aire géographique. Ainsi, les phylotypes rares archéens représenteraient la majorité de la diversité et selon Galand *et al.* (2009a) auraient, dans les milieux aquatiques, des profils de distribution endémiques à certaines masses d'eau. Leur distribution serait donc conditionnée par des mécanismes de spéciation, d'extinction, de dispersion ou d'interactions, équivalents à ceux régissant les phylotypes abondants (Martiny *et al.*, 2006).

L'activité des microorganismes rares est encore mal connue, pourtant Campbell *et al.* (2011) postulent que certains *taxa* rares restent toujours rares mais contrairement aux idées reçues, ne seraient pas dormants, mais actifs. Ces microorganismes pourraient être impliqués dans des fonctions physiologiques clés, mais également maintenir un équilibre permettant à la communauté de s'adapter rapidement en réponse à des changements environnementaux, ou encore constituer un réservoir de gènes permettant des adaptations évolutives. Plusieurs

raisons sont évoquées pour le maintien en faible abondance de ces *taxa* comme une faible disponibilité des ressources dans le milieu, et/ou une faible compétitivité par rapport aux ressources et/ou une susceptibilité à la lyse virale. Ceci pourrait également résulter d'une stratégie adaptative pour échapper aux phénomènes de prédation (Pedros-Alio, 2006).

2.2.3 Distribution des *Euryarchaeota*

Le nombre croissant de séquences environnementales affiliées aux *Euryarchaeota* permet aujourd'hui d'attester de leur ubiquité dans les écosystèmes aquatiques. Elles ont initialement été décrites comme faisant partie des Marine Group II, III affiliées aux *Thermoplasmatales*. La plupart des études conduites en milieu marin a par ailleurs permis de montrer la prédominance du MGII dans les eaux de surface (Auguet *et al.*, 2009; DeLong, 1998). De plus, une étude conduite en mer Méditerranée a permis de montrer qu'il existait deux clusters différents de MGII dont l'abondance était liée à des saisons différentes (Galand *et al.*, 2010).

D'autres lignées, moins décrites dans la littérature, sont souvent retrouvées dans les eaux intérieures (inland water) et sont essentiellement membres de deux clusters : le LDS (Lake Dagow Sediment, (Glissman *et al.*, 2004)) et le RC-V (Rice Cluster-V, (Großkopf *et al.*, 1998)). Ces lignées sont classiquement retrouvées dans des rivières (Galand *et al.*, 2008; Herfort *et al.*, 2009) ou encore des lacs (Auguet and Casamayor, 2008; Lliros *et al.*, 2008) mais semblent être ubiquitaires et très diversifiées dans les écosystèmes aquatiques d'eau douce en général, suggérant leur rôle clé dans ces milieux (Auguet *et al.*, 2009).

Enfin, les *Euryarchaeota* méthanogènes prospèrent dans les environnements anoxiques riches en matière organique (sédiments marins ou lacustres, sources chaudes...) et leur rôle primordial dans le cycle du carbone s'illustre par l'estimation à un milliard de tonnes de méthane produit du fait de leur activité (Thauer, 1998).

2.2.4 Distribution des *Thaumarchaeota*

Comme décrit précédemment, il existe plusieurs groupes de *Thaumarchaeota* et leur différenciation phylogénétique semble conditionnée par le milieu qu'ils occupent. En effet, la majorité des séquences affiliées au groupe 1.1a (MGI) sont retrouvées en milieu aquatique alors que le groupe 1.1b semble caractéristique des milieux terrestres (Ochsenreiter *et al.*,

2003). Dans les écosystèmes marins, les *Thaumarchaeota* du groupe 1.1a sont abondantes dans les zones méso- et bathypélagiques (Karner *et al.*, 2001). En milieu lacustre, ces microorganismes sont abondants dans le neuston (Auguet and Casamayor, 2008), mais aussi en surface et au niveau de l'oxycline où les apports d'oxygène par les eaux de surface et l'apport en ammonium par les couches profondes rassemblent des conditions favorables pour la réalisation du processus d'oxydation de l'ammonium (Lliros *et al.*, 2010; Pouliot *et al.*, 2009). Dans certains milieux aquatiques, on retrouve également des *Thaumarchaeota* affiliées au groupe SAGMCG-1 (Auguet *et al.*, 2012) initialement décrit dans des mines Sud-Africaines (Takai *et al.*, 2001b). Ce groupe-sœur du 1.1a n'a pas été détecté dans les écosystèmes marins, et est retrouvé principalement dans les milieux d'eau douce, où sa fréquence et son abondance relative importantes (Auguet *et al.*, 2012) suggèrent un rôle dans le fonctionnement de ces écosystèmes.

Il a été mis en évidence une différenciation verticale de ces deux groupes en milieu lacustre, où SAGMCG-1 est associé aux eaux de surface tandis que le groupe 1.1a est retrouvé préférentiellement dans les couches plus profondes (Auguet *et al.*, 2012). En outre, l'étude récente conduite par Auguet et Casamayor (2013) dans des lacs de haute montagne a permis de montrer que le groupe 1.1a était retrouvé dans les écosystèmes caractérisés par de plus faibles concentrations en nitrite et carbone organique dissous alors que le groupe SAGMCG-1 serait préférentiellement retrouvé dans des lacs présentant de faibles pH (< 5) et des concentrations en nitrites supérieures à 0.12 μM (Auguet and Casamayor, 2013). Les facteurs influençant la structuration de ces communautés sont à ce jour encore mal connues même si l'on peut présumer d'un effet du pH ou encore des phénomènes de photo-inhibition démontrés récemment sur les souches de *N. maritimus* et *N. devanattera* (Merbt *et al.*, 2012). Ces résultats sont néanmoins en contradiction avec la forte abondance de *Thaumarchaeota* du groupe 1.1a dans le neuston de lacs de haute montagne (Auguet and Casamayor, 2008). L'ensemble de ces observations suggère une sélection des différents groupes par des conditions environnementales sélectives offrant des niches écologiques particulières, pouvant expliquer en partie le succès écologique des *Archaea* planctoniques dans ces milieux.

Cependant, la diversité des *Thaumarchaeota* est certainement bien plus complexe que celle décrite. En effet, il a été mis en évidence une microdiversité au sein du groupe 1.1a, pour lequel différents clusters ont été décrits en milieu marin, en lien avec des profondeurs et des localisations géographiques différentes (Garcia-Martinez and Rodriguez-Valera, 2000; Sintes *et al.*, 2013). Dans le même esprit, une étude récente de métagénomique menée dans les eaux côtières de surface de l'océan Atlantique a permis de suggérer la présence de deux

populations divergentes de *Thaumarchaeota* du groupe 1.1a (Tully *et al.*, 2012). La présence de deux clusters initialement décrits dans différents milieux marins (Massana *et al.*, 2000), au sein de ce groupe a été confirmée lors d'une étude conduite en mer Méditerranée (Galand *et al.*, 2010). Ces microorganismes, distincts phylogénétiquement, pourraient donc occuper des fonctions écosystémiques différentes au sein de niches écologiques distinctes et correspondre à des écotypes (Koeppel *et al.*, 2008).

Toutes ces études laissent entrevoir l'hypothèse selon laquelle les communautés d'*Archaea* seraient constituées d'entités phylogénétiques distinctes, ou écotypes, c'est-à-dire de clusters adaptés à des conditions environnementales particulières (profondeur, nutriments...) suite à des adaptations métaboliques. Cependant, les *Euryarchaeota* mésophiles (pour lesquelles, excepté les méthanogènes et les halophiles, aucun représentant n'est cultivé) ont fait l'objet de peu de travaux, contrairement aux *Thaumarchaeota*. Ceci s'explique en partie par l'existence de représentants cultivés de ce groupe mais également par leur implication maintenant reconnue dans un cycle biogéochimique majeur : le cycle de l'azote.

2.3 Implication des *Archaea* dans le cycle de l'azote

Pour croître, se reproduire et maintenir leur intégrité, les êtres vivants puisent de façon sélective dans leur environnement un nombre limité d'éléments chimiques qui s'agencent en biomolécules organiques, unités de base des cellules. Ainsi, l'eau, l'oxygène, le carbone, l'azote et l'hydrogène constituent 95% de la composition atomique des organismes tandis que les 5% restants sont composés de phosphore, soufre, chlore, sodium, potassium, calcium, magnésium et autres éléments trace (aluminium, fer, zinc...). Par ailleurs, la circulation et le recyclage continu des éléments sont assurés au sein de réseaux trophiques via des flux de matière entre les organismes autotrophes, hétérotrophes et l'environnement. Par conséquent, le bon déroulement des cycles de matière est assuré par une complémentarité écophysiologique des différents intervenants. Ainsi, la pérennité des écosystèmes est assurée par le maintien d'une autorégulation ou homéostasie. Les microorganismes jouent un rôle primordial dans le bon déroulement de ces mécanismes par leur implication dans des fonctions très spécifiques. C'est notamment le cas de certaines *Archaea* dont le rôle biogéochimique dans le cycle de l'azote, et plus particulièrement dans l'oxydation de l'ammonium, est aujourd'hui confirmé dans de nombreux milieux.

2.3.1 L'azote, un composé essentiel dans les écosystèmes

L'azote est l'élément le plus abondant de l'atmosphère (78% en volume). Il s'agit d'un élément nutritif essentiel pour les plantes et les animaux, représentant environ 5% de leur poids sec et entrant dans la composition des acides aminés, des acides nucléiques et de nombreuses autres biomolécules. Un grand nombre et une importante diversité de microorganismes interviennent dans le cycle de l'azote et ce, de façon hiérarchisée (Ramade, 2009). Comme tous les éléments chimiques, l'azote subit, au sein des écosystèmes, une série de transformations biologiques complexes. Dans les milieux aquatiques, 5 processus de transformation majeurs de cet élément se produisent : fixation, assimilation, minéralisation, nitrification, dénitrification (Figure 6). Son cycle biogéochimique revêt donc un intérêt particulier puisqu'il va réguler la quantité d'azote disponible pour le réseau trophique.

Il est à noter que la nitrification est un processus se déroulant en deux étapes au cours desquelles l'ammonium (NH_4^+) est oxydé en nitrite (NO_2^-) puis en nitrate (NO_3^-). La phase d'oxydation de l'ammonium est la première étape ; elle repose sur l'activité d'un complexe enzymatique, l'ammonium monooxygénase (AMO), constitué de trois sous unités A, B et C, codées respectivement par les gènes *amoA*, *amoB* et *amoC*.

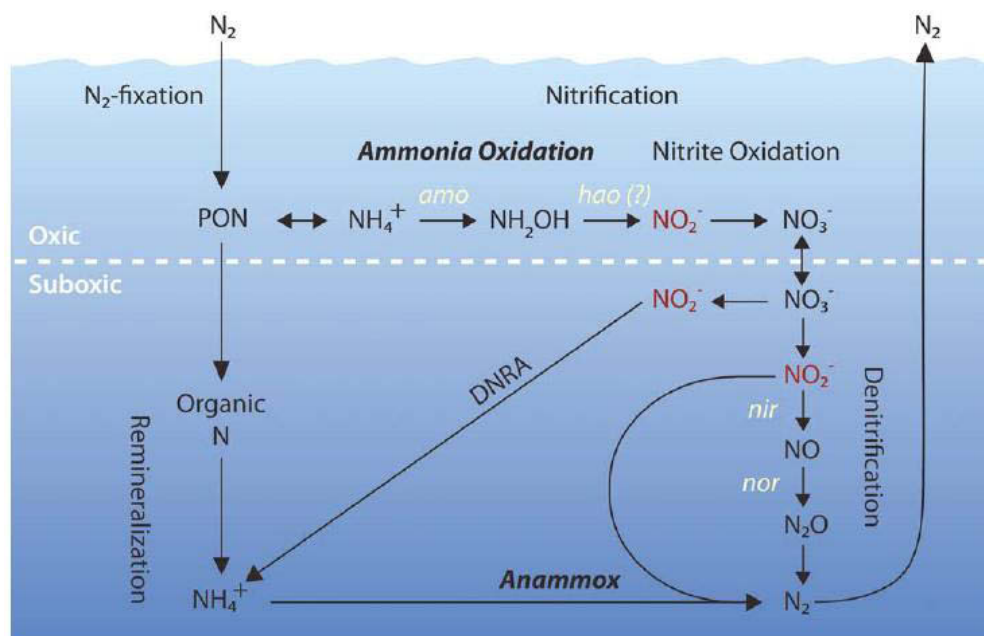


Figure 6 : Schématisation des différentes étapes de transformations des composés azotés par les microorganismes dans les milieux marins, réalisées dans la zone oxic, anoxique et à l'interface oxic/anoxique (basé en partie sur (Arrigo, 2005)). Les nitrites sont représentés en rouge pour souligner le rôle central de cet intermédiaire métabolique. Les gènes fonctionnels impliqués dans la nitrification sont : *amo*, ammonia monooxygenase ; *hao*, bacterial hydroxylamine oxidoreductase. Ceux impliqués dans la dénitrification sont : *nir*, nitrite reductase et *nor*, nitric oxide reductase. PON : azote particulaire organique ; DNRA : réduction dissimilatrice des nitrates, d'après Francis *et al.* (2005).

2.3.2 Diversité des *Archaea* oxydant l'ammonium (AOA)

Comme indiqué dans le paragraphe précédent, parmi les différents processus impliqués dans les transformations de l'azote, les *Archaea* semblent jouer un rôle primordial dans la première étape de la nitrification, à savoir l'oxydation de l'ammonium en nitrite. Les gènes codant les enzymes impliquées dans cette étape (*amo*) sont utilisés comme marqueurs phylogénétiques, au même titre que les gènes d'ADNr 16S.

Alors que la capacité des bactéries à oxyder l'ammonium est connue depuis plus d'une centaine d'années (Kowalchuk and Stephen, 2001), celle de certaines *Archaea* à réaliser ce processus n'a été révélée que très récemment grâce à l'isolement d'un clone dérivé d'une banque de fosmides (Figure 7). Ce clone contenait un fragment d'ADN génomique porteur à la fois des gènes fonctionnels *amoA* et *amoB* et des gènes codant les ARNr 16S et 23S (Treusch *et al.*, 2004; Treusch *et al.*, 2005). L'affiliation taxonomique des gènes d'ADNr 16S a montré qu'ils s'intégraient au niveau du cluster 1.1b des *Thaumarchaeota*, ceci constituant le premier indice suggérant une possible implication des *Archaea* dans l'oxydation de l'ammonium. En lien avec le processus de nitrification, le fragment d'ADNg portait également un gène homologue à celui codant une nitrite réductase (*nirK*), enzyme impliquée dans la dénitrification.

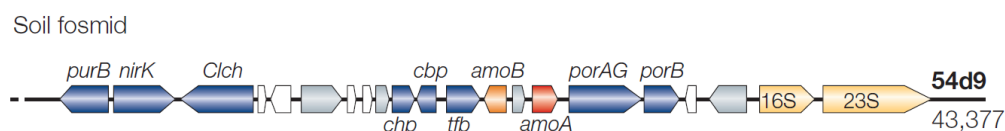


Figure 7 : Représentation schématique du fragment génomique porté par le fosmide 54d9 (43 kpb) et affilié aux *Thaumarchaeota*. Les gènes codant les ARNr 16S et 23S sont représentés en jaune, les gènes avec des fonctions assignées en bleu, les gènes hypothétiques conservés en gris et les gènes hypothétiques en blanc. Les gènes *amo* sont représentés en orange et rouge, d'après Treusch *et al.* (2005).

Actuellement, l'existence d'un processus de nitrification chez les *Archaea* est un fait reconnu, les gènes d'ADNr 16S et *amoA* d'*Archaea* étant retrouvés dans de nombreux milieux à travers le monde. Les séquences isolées depuis cette découverte indiquent que seules les populations affiliées aux *Thaumarchaeota* possèdent les gènes *amo*. Actuellement, plusieurs souches appartenant à ce *phylum* et pour lesquelles la capacité à réaliser la nitrification est avérée, ont été isolées en culture pure ou dans des enrichissements (Table 1).

Table 1 : Bilan des caractéristiques physiologiques et génomiques des souches de *Thaumarchaeota* oxydant l'ammonium, isolées en culture pure ou en culture d'enrichissement. ND : non déterminé.

Groupe	1.1a				SAGMCG-1	1.1b		ThAOA HWCIII
Souche	<i>Nitrosopumilus maritimus SCM1</i>	<i>Cenarchaeum symbiosum A</i>	<i>Nitrosoarchaeum</i>		<i>Nitrosotalea devanattera</i>	<i>Nitrososphaera</i>		<i>Nitrosocaldus yellowstonii</i>
			<i>limnia SFB1</i>	<i>koreensis MY1</i>		<i>viennensis EN76</i>	<i>gargensis</i>	
Habitat	Eau de mer tropicale	Symbionte	Sédiment	Sol			Source chaude	Sédiment
Culture	Pure	Coculture <i>Axinella mexicana</i>	Enrichissement					
Température de croissance (°C)	28	10	22	25	25	35	46	72
pH Optimal	7-7,2	ND	7-7,2	6-8	4,5	7,5	7-7,4	8
Taux de croissance (d ⁻¹)	0,65 à 30°C	ND	0,2	ND	0,37	0,043 (auto) 0,53 (mixo)	ND	0,8
Taille (µm)	0,2-0,9	0,5	0,2	0,3-0,5	0,5-0,8	0,5-0,8	0,9	0,5
[NH ₄ ⁺] Inhibitrice (mM)	2-3	ND	ND	20	< 50	10-15	< 3	ND
[NH ₃] Inhibitrice (mM)	18 à 27.10 ⁻³ (pH7)	ND	ND	145.10 ⁻³ (pH7)	< 9.10 ⁻³	0,51-0,75	< 88.10 ⁻³ (pH7)	ND
Taille du génome (Mb)	1,65	2,05	1,77	1,6	ND	ND	> 2,83	1,43
% de séquences codantes	91,9	91,2	83	ND	ND	ND	ND	ND
% en GC	34,2	57,7	32	32,7	ND	ND	48	ND
Nombre de gènes prédits	1840	2066	2171	1317	ND	ND	> 3200	ND
Opéron 16S-23S	1				ND	ND	1	ND
Opéron 5S					ND	ND		ND
tRNAs	44	45	45	42	ND	ND	ND	ND
Gènes <i>ureC</i>	Non	Oui	ND	Non	ND	Oui	ND	ND
Référence	Konneke <i>et al.</i> , 2005 Martens Habbena <i>et al.</i> , 2009	Preston <i>et al.</i> , 1996 Hallam <i>et al.</i> , 2006	Blainey <i>et al.</i> , 2011 Mosier <i>et al.</i> , 2012	Kim <i>et al.</i> , 2011 Jung <i>et al.</i> , 2011	Lethovirta <i>et al.</i> , 2011	Tourna <i>et al.</i> , 2011	Hatzenpichler <i>et al.</i> , 2008	de la Torre <i>et al.</i> , 2008

En milieu naturel, Treusch *et al.* (2005) ont montré une augmentation de la transcription des gènes *amoA* d'*Archaea* en relation avec les concentrations en ammonium. Ce marqueur fonctionnel, dont la séquence est différente de celle retrouvée chez les bactéries, est désormais couramment utilisé pour établir des liens phylogénétiques au sein des AOA (Junier *et al.*, 2010). L'augmentation de la quantité de données de métagénomique, que ce soit pour des

environnements aquatiques ou terrestres (Treusch *et al.*, 2005; Venter *et al.*, 2004) permet de disposer d'un nombre de séquences de gènes *amoA* archéens de plus en plus important. Leur analyse et leur alignement ont révélé la présence de régions conservées propices à la détermination d'amorces, indispensables pour des approches PCR (Schleper and Nicol, 2010). L'utilisation de ces amorces pour l'étude ciblée d'environnements aquatiques a permis d'enrichir à leur tour les bases de données en séquences *amoA* provenant d'une grande diversité d'environnements et de confirmer la large distribution des AOA (Auguet *et al.*, 2011).

Pour les *Archaea*, la phylogénie construite sur la base du gène *amoA* est cohérente avec celle construite à l'aide des gènes d'ADNr 16S (Figure 3). L'arbre phylogénétique réalisé à partir de ce gène (méta-analyse de l'ensemble des séquences *amoA* disponibles dans les bases de données) montre un découpage en cinq clusters (Figure 8) (Pester *et al.*, 2011).

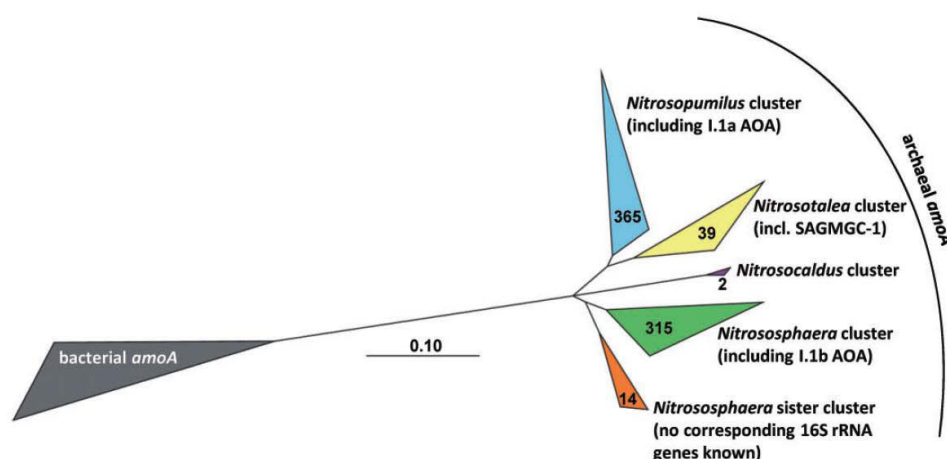


Figure 8 : Arbre consensus déterminé à partir de l'ensemble des séquences *amoA* d'*Archaea* disponibles dans les bases de données en comparaison des séquences ADNr 16S et illustrant la présence de 5 clusters majeurs. Le nom couramment utilisé dans les phylogénies ADNr 16S pour désigner chaque groupe est indiqué entre parenthèses, d'après Pester *et al.* (2011a).

Dans un but d'harmonisation, une nouvelle nomenclature a été proposée pour ces clusters *amoA* en lien avec les groupes ADNr 16S auxquels ils se rattachent. Ces clusters sont : le cluster *Nitrosopumilus* caractérisant le groupe 1.1a en ADNr 16S, le cluster *Nitrosotalea* correspondant au groupe SAGMCG-1, le cluster *Nitrososphaera* correspondant au groupe 1.1b, le cluster *Nitrosocaldus* correspondant au groupe HWCGIII et le groupe sœur de *Nitrososphaera* pour lequel aucune correspondance ADNr 16S n'a été établie dû à l'absence de représentant cultivé pour ce groupe.

De plus, l'étude conduite par Francis *et al.* (2005) visant à évaluer l'ubiquité et la diversité des AOA dans une colonne d'eau marine a permis de définir la présence de deux groupes de gènes *amoA* phylogénétiquement différents, nommés A et B. Par la suite, ces groupes ont été décrits comme étant deux écotypes distincts, l'un retrouvé dans les eaux de surface alors que l'autre serait présent dans les eaux plus profondes (750 m, (Hallam *et al.*, 2006b; Suzuki *et al.*, 2004)). La différenciation phylogénétique et spatiale de ces écotypes a été confirmée par d'autres études et il a été suggéré que leur distribution pourrait être conditionnée par des phénomènes de sensibilité à la photo-inhibition ou encore par un manque de compétitivité pour l'ammonium vis-à-vis du phytoplancton en présence de lumière (Beman *et al.*, 2008; Mincer *et al.*, 2007).

De la même façon en milieu lacustre, les conditions environnementales conditionneraient la diversité des AOA et leur distribution saisonnière. Ainsi, Pouliot *et al.* (2009) ont montré dans lac Arctique particulier (salinité élevée) que les AOA sont présentes dans la zone euphotique lacustre, avec néanmoins un maximum d'abondance au niveau de l'interface oxygène-anoxique (ou oxycline). Dans des lacs de haute montagne oligotrophes, Auguet *et al.* (2011, 2012) ont démontré des variations saisonnières et verticales des assemblages d'AOA en lien avec des concentrations variables en nitrite et en ammonium. L'ensemble de ces résultats suggère donc la présence d'écotypes au sein des AOA.

2.4 Métabolismes des *Archaea* nitrifiantes

L'isolement d'AOA en culture pure ou au sein d'enrichissements a permis d'accéder à une partie des informations concernant leur physiologie. Ainsi, il a été suggéré que ces microorganismes tiraient leur énergie de l'oxydation de l'ammonium en nitrite.

2.4.1 Eléments intervenant dans la nitrification chez les AOA

L'enzyme clé de la nitrification, pour les AOB comme pour les AOA, est l'ammonium monooxygénase (AMO). Cette enzyme est formée d'un complexe constitué de 3 sous-unités A, B et C. Chez les AOB l'ordre des gènes codant ces sous-unités, dans l'opéron, est généralement BCA (Bothe *et al.*, 2000; Nicol and Schleper, 2006), ce qui n'est pas le cas

chez les AOA, pour lesquelles cet ordre varie d'une lignée à l'autre (Bartossek *et al.*, 2012; Blainey *et al.*, 2011; Kim *et al.*, 2011a; Spang *et al.*, 2010; Walker *et al.*, 2010).

Les complexes AMO bactérien et archéen sont très ressemblants, pourtant, les études tendent à penser que le mécanisme de nitrification serait différent (Klotz, 2011). Chez les bactéries, l'oxydation de l'ammonium se déroule en deux étapes. L'oxydation de l'ammonium en hydroxylamine (NH_2OH) est réalisée grâce à l'AMO, puis l'hydroxylamine oxydoréductase (HAO) permet d'oxyder le NH_2OH (Figure 9.A). Pour les *Archaea*, les étapes de cette réaction sont toujours incertaines et plusieurs processus sont évoqués (Figure 9.B). La première hypothèse propose une voie similaire à celle des AOB. L'oxydation de l'ammonium en nitrite se ferait via le NH_2OH mais les enzymes impliquées dans la fin de la réaction (HAO, cytochromes) seraient différentes. Actuellement, aucun homologue pour les gènes codant l'HAO et les cytochromes n'a encore été mis en évidence chez les *Archaea* (Simon and Klotz, 2012; Walker *et al.*, 2010). D'autres travaux tendent à montrer l'existence chez les AOA d'une voie métabolique fondamentalement différente de la première. Dans cette nouvelle voie, l'intermédiaire ne serait plus le NH_2OH mais le nitroxyle (HNO). Ainsi, l'enzyme intervenant dans la seconde étape et responsable de l'oxydation du HNO en nitrite serait la nitroxyle oxydoréductase (NXOR) (Schleper and Nicol, 2010; Walker *et al.*, 2010). Dans ce modèle, le monoxyde d'azote (NO), produit après réduction des nitrites, servirait d'activateur de l'AMO. Le métabolite final libéré serait alors le diazote (N_2) contrairement aux AOB qui libèrent du N_2O , un puissant gaz à effet de serre (Schleper and Nicol, 2010).

L'analyse des enzymes intervenant dans la nitrification chez les bactéries et les *Archaea* montre qu'il existe d'autres éléments les différenciant comme notamment leur affinité vis à vis des cofacteurs. Ainsi, alors que les enzymes bactériennes sont généralement fer-dépendantes, la plupart des enzymes archéennes sont cuivre-dépendantes (Glass and Orphan, 2012; Godfrey and Glass, 2011). Cette différence pourrait en partie expliquer les différences d'abondance observées pour ces deux groupes en fonction des environnements (Urakawa *et al.*, 2011). Ainsi, les milieux marins plus riches en cuivre, pourraient favoriser la croissance des AOA.

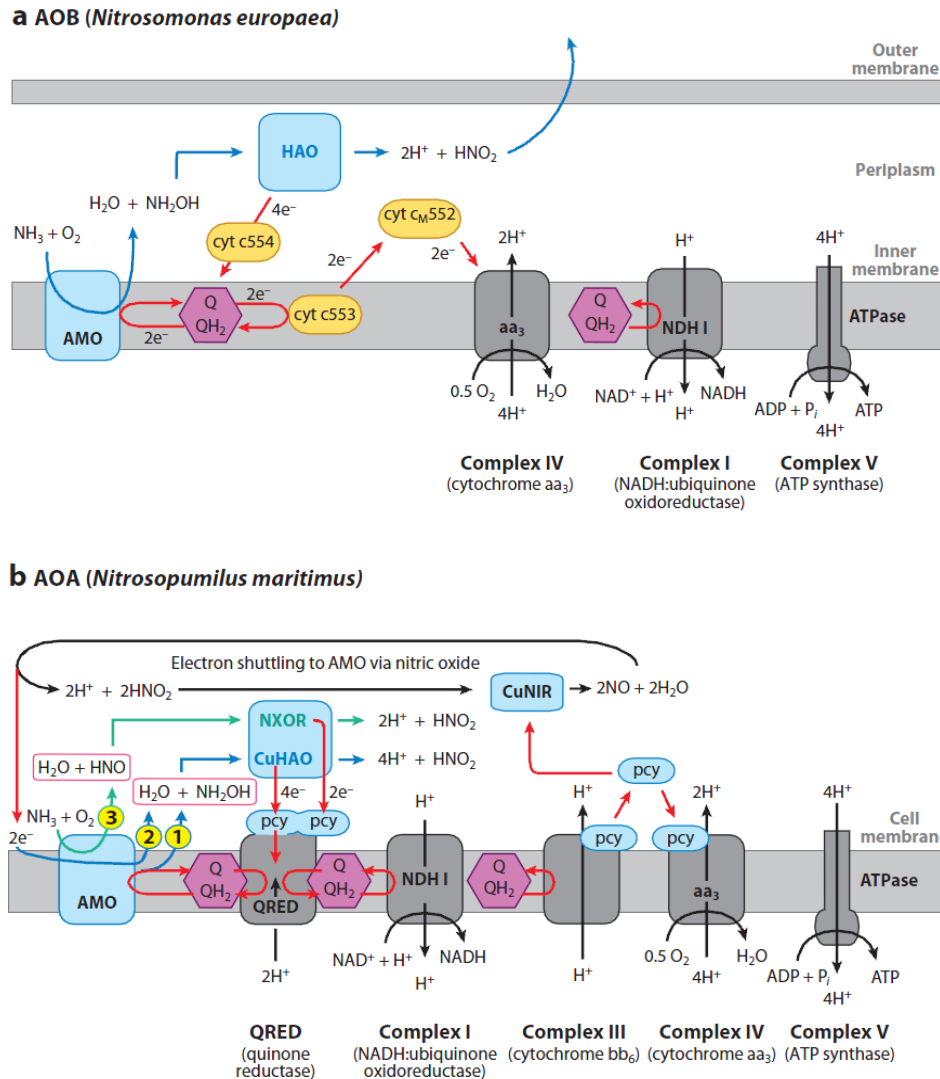


Figure 9 : (A) Représentation de la cascade métabolique responsable de la nitrification chez les AOB. En présence d'oxygène, l'ammonium est oxydé par l'ammonium monooxygénase (AMO) en hydroxylamine (NH_2OH), lui aussi à son tour oxydé en nitrite par le module N-oxyde-Ubiquinone Redox (NURM). Ce module consiste en une hydroxylamine oxydoréductase (HAO), un cytochrome c554 et une quinone réductase cM552, d'après Stahl et de la Torre (2012).

(B) Représentation de la cascade métabolique responsable de la nitrification chez les AOA. Trois mécanismes sont représentés : dans les mécanismes 1 et 2, le produit de l'oxydation de l'ammonium serait l'hydroxylamine. Ces deux mécanismes diffèrent par l'origine des électrons nécessaires pour initier l'oxydation de l'ammonium par l'AMO. Au cours du mécanisme 1, proche de celui retrouvé chez les AOB, les électrons produits par l'oxydation de l'hydroxylamine en nitrites par une CuHAO sont transférés à des transporteurs pcy puis à des quinones. Deux électrons sont alors recyclés vers l'AMO et deux autres sont transférés vers la chaîne de transport des électrons. Au cours du mécanisme 2, le NO produit par la réduction des nitrites par une CuNIR serait la source d'électrons pour l'AMO. La possibilité que le HNO soit produit sous l'action de l'AMO est représenté dans la voie 3. Les flèches rouges illustrent le flux électronique, les cadres bleus des protéines contenant du cuivre, les hexagones avec des Q, des quinones oxydées et ceux avec des QH_2 , des quinones réduites.

Abréviations: HAO, hydroxylamine oxydoréductase ; NO, nitric oxide ; HNO, nitroxyde ; CuHAO, cuivre hydroxylamine oxydoréductase ; CuNIR, nitrite réductase dépendante du cuivre ; NXOR, nitroxyde oxydoréductase putative ; pcy, plastocyanines ; NH_2OH , hydroxylamine; QRED, quinone réductase, d'après Stahl et de la Torre (2012).

2.4.2 Assimilation autotrophe du carbone

Les microorganismes nitrifiants autotrophes tirent leur énergie de l'oxydation de l'ammonium et utilisent le CO₂ comme seule source de carbone pour la synthèse de leur biomasse. Qu'ils soient bactériens ou archéens, ces microorganismes chimiolithoautotrophes sont très difficiles à faire pousser et/ou à maintenir en culture. Généralement, l'assimilation du CO₂ se fait via le cycle de Calvin-Bassham-Benson, qui est initié par une enzyme clé, la ribulose 1,5-bisphosphate carboxylase/oxygénase (RubisCO). La plupart des algues et de nombreuses bactéries, parmi lesquelles les AOB (Arp *et al.*, 2007), réalisent ce processus. Cependant, cinq voies métaboliques alternatives ont également été décrites pour la fixation du CO₂ chez les procaryotes mais elles ne comptent que pour une faible part de la fixation totale du carbone (Menendez *et al.*, 1999). Il s'agit du cycle des acides tricarboxyliques (ou cycle de Krebs), du cycle de l'acétyl-Co-A, du cycle du dicarboxylate-hydroxybutyrate, du 3-hydroxypropionate bicycle décrit chez les *Chloroflexaceae*, et d'une version modifiée de ce dernier, le cycle du 3-hydroxypropionate 4-hydroxybutyrate.

Les travaux portant sur l'étude de l'activité *in situ* des AOA montrent un lien étroit entre les cycles du carbone et de l'azote chez ces organismes (Agogué *et al.*, 2008; Beman *et al.*, 2008; Santoro *et al.*, 2010; Stewart *et al.*, 2012; Yakimov *et al.*, 2011). Cependant, l'assimilation du carbone se ferait différemment de ce qui a été décrit jusqu'à maintenant. En effet, le mécanisme proposé correspondrait à une modification du cycle 3-hydroxypropionate alors appelé 3-hydroxypropionate 4-hydroxybutyrate ou 3HP/4HB (Figure 10). Ce processus de fixation du carbone inorganique a été suggéré pour les espèces *N. maritimus*, *C. symbiosum* et des AOA planctoniques marines sur la base de données de génomique (Hallam *et al.*, 2006b; Martin-Cuadrado *et al.*, 2008; Walker *et al.*, 2010). Malgré une conservation de certains intermédiaires et de certaines réactions avec le cycle du 3-hydroxypropionate (Offre *et al.*, 2010; Yakimov *et al.*, 2011), le cycle identifié chez les AOA aurait subi une évolution différente (Berg *et al.*, 2010; Blainey *et al.*, 2011; Hallam *et al.*, 2006b; Park *et al.*, 2010; Tourna *et al.*, 2011; Walker *et al.*, 2010). L'utilisation d'une voie préférentielle pour la fixation du carbone a d'importantes conséquences sur l'adaptation écologique de ces microorganismes (Berg *et al.*, 2010). Dans les cycles 3HP/4HB et 3-hydroxypropionate, c'est le bicarbonate (HCO₃⁻) qui est assimilé alors que dans celui de Calvin, c'est le CO₂. Les AOA pourraient ainsi privilégier le cycle 3HP/4HB dans les écosystèmes marins dans la mesure où l'ion HCO₃⁻ est la forme carbonée la plus fréquemment rencontrée.

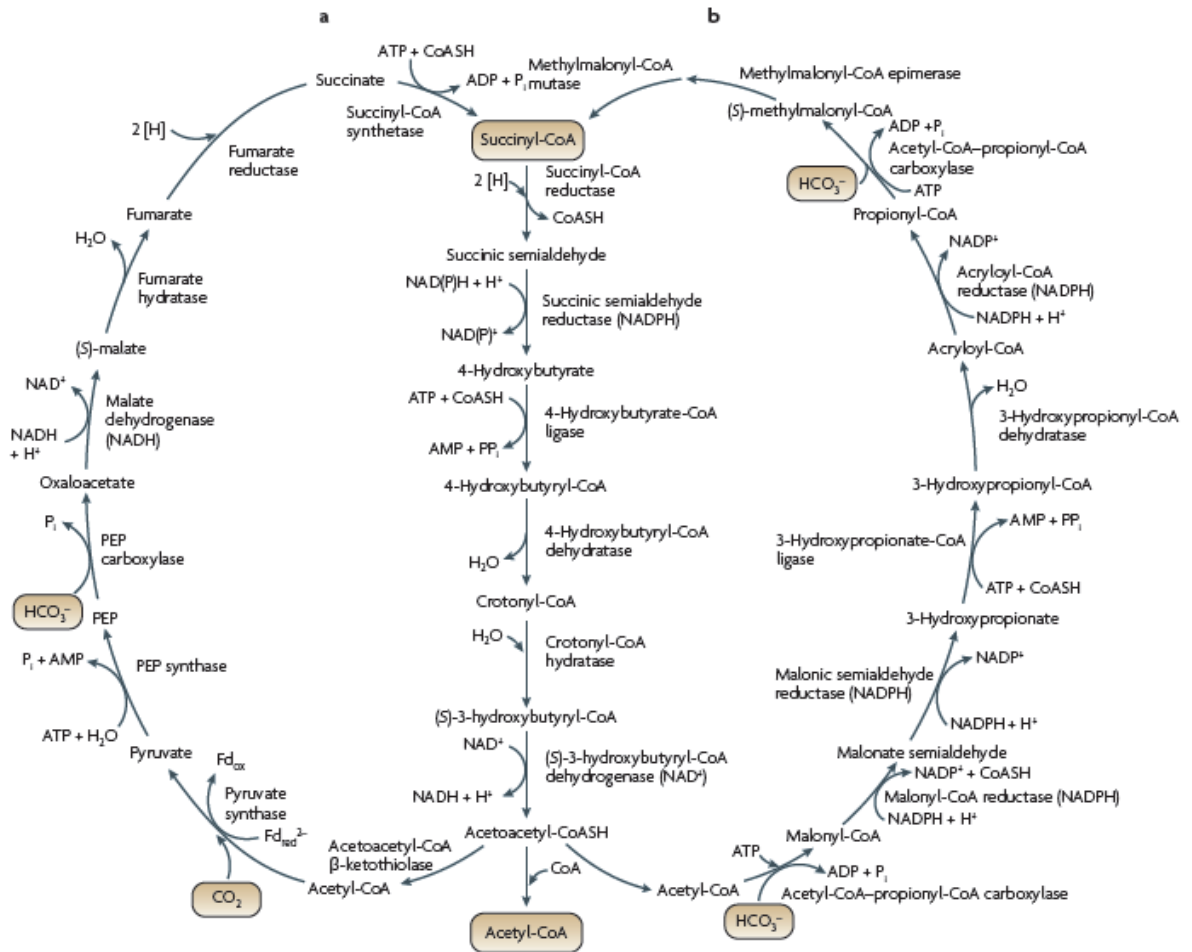


Figure 10 : Voies d'assimilation du CO₂ chez les *Thaumarchaeota* autotrophes. (a) Cycle dicarboxylate-hydroxybutyrate retrouvé chez les *Desulfurococcales* et *Thermoproteales* et (b) cycle hydroxypropionate-hydroxybutyrate (3HP/4HB) retrouvé chez les *Sulfolobales* et les *Thaumarchaeota*, d'après Berg *et al.* (2010).

Malgré la détection dans la quasi-totalité des milieux de la présence de gènes *amoA* archéens et la démonstration d'un métabolisme de type chimolithotrophie chez certaines espèces cultivées, il semblerait que tous les mécanismes à l'origine des sources de carbone et d'énergie chez les *Thaumarchaeota* ne soient pas encore complètement décryptés. Cette hypothèse est le résultat de travaux menés sur des milieux particuliers comme les océans polaires, les océans profonds ou encore certains sols. Dans ces milieux, différents auteurs ont observé, en plus des gènes *amoA*, la présence en abondance de gènes d'uréases d'*Archaea* (*ureC*), impliqués dans la dégradation de l'urée (Baker *et al.*, 2012; Konstantinidis *et al.*, 2009; Yakimov *et al.*, 2011). Ils suggèrent que la dégradation de l'urée fournirait aux *Thaumarchaeota* une source d'énergie via la production d'ammonium et une source de carbone *via* la production de CO₂ (Figure 11). Cependant, ce mécanisme étant

énergétiquement coûteux (synthèse des uréases et transport de l'urée), il ne serait utilisé que dans les milieux où la disponibilité de l'ammonium est faible ou intermittente.

Chez les *Thaumarchaeota*, la première démonstration de la présence d'uréases au niveau du génome a été faite pour l'espèce symbiotique *C. symbiosum*. Ces gènes lui confèrent la capacité d'utiliser l'urée rejetée par son hôte comme source de composés nécessaires à la nitrification et donc à sa croissance (Hallam et al., 2006b). En revanche, les microorganismes nitrifiants comme *N. maritimus* (Walker et al., 2010) et *N. limnia* (Blainey et al., 2011) en sont dépourvus.

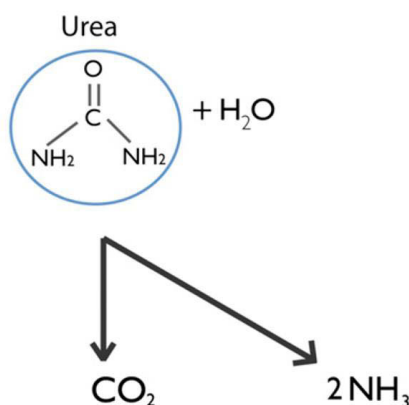


Figure 11 : Mécanisme d'uréolyse proposé pour les *Thaumarchaeota*, comme source de carbone et d'énergie, d'après Alonso-Saez *et al.* (2012).

2.4.3 Diversité des métabolismes retrouvés chez les AOA

Les *Thaumarchaeota* nitrifiantes sont considérées comme des chimiolithoautotrophes (Francis *et al.*, 2005; Venter *et al.*, 2004) incorporant du carbone inorganique dissous et utilisant l'ammonium comme source d'énergie. Cependant, leur capacité à incorporer d'autres substrats indique des voies métaboliques certainement plus diversifiées (Herndl *et al.*, 2005; Ingalls *et al.*, 2006).

En effet, des mesures d'activité cellulaire, conduites sur les membres du groupe 1.1a, ont montré qu'elles pouvaient incorporer divers composés organiques (Agogué *et al.*, 2008; Ingalls *et al.*, 2006; Varela *et al.*, 2011) parmi lesquels des acides aminés (Ouverney and Fuhrman, 2000). Ces résultats ont par la suite été confirmés par des études conduites à plus grande échelle dans l'océan Atlantique (Herndl *et al.*, 2005; Teira *et al.*, 2004). Cependant, la mise en évidence de la capacité de certaines espèces considérées autotrophes à incorporer de

composés organiques suggère une croissance mixotrophe pour certaines d'entre elles (Hallam *et al.*, 2006a; Hansman *et al.*, 2009; Ingalls *et al.*, 2006). Ceci a été montré notamment lors de l'annotation du génome de *N. maritimus* révélant la présence de transporteurs de molécules organiques, puis confirmé expérimentalement (Walker *et al.*, 2010). Ainsi cette espèce serait capable de métaboliser des molécules organiques comme le pyruvate et l'alpha-kétoglutarate alors qu'elle est incapable d'utiliser les acides gras et les acides aminés (Stahl and de la Torre, 2012).

Dans le même esprit, des expérimentations conduites en bioréacteurs suggèrent que certaines *Thaumarchaeota* proches du groupe 1.1b portant les gènes *amoA*, seraient incapables d'assimiler le CO₂ et d'oxyder l'ammonium (Mussmann *et al.*, 2011). Les gènes *amoA* portés par ces populations occuperaient donc une fonction différente comme par exemple la dégradation de composés organiques. Plus largement, ceci permet de s'interroger sur la réelle capacité des *Thaumarchaeota* à réaliser la nitrification dans tous les écosystèmes aquatiques et terrestres où la présence de gènes *amoA* a été détectée.

2.5 Facteurs contrôlant l'abondance et la répartition des AOA en milieu aquatique

L'abondance et la large distribution des AOA dans les écosystèmes posent bon nombre de questions quant aux réelles limites de leurs capacités d'adaptation. Ainsi, plusieurs facteurs guideraient leur aptitude à pouvoir coloniser une niche écologique parmi lesquels la température (l'expression des gènes *amoA* serait favorisée à 30°C (Nicol *et al.*, 2008; Tourna *et al.*, 2008)), ou encore de faibles teneurs en oxygène (<10 µM (Beman *et al.*, 2008; Molina *et al.*, 2010)). Parmi les facteurs les plus importants, le statut trophique du milieu, la concentration en substrat, mais aussi le pH ou encore la salinité joueraient un rôle majeur (Erguder *et al.*, 2009).

2.5.1 Influence de la concentration en ammonium

L'un des principaux facteurs influençant l'établissement des communautés d'AOA serait la concentration en ammonium. Ceci a été montré avec l'espèce *N. maritimus* qui, contrairement aux AOB, présente une très grande affinité pour l'ammonium ($K_m = 133$ nM de NH₄⁺) (Martens-Habbena *et al.*, 2009; Martens-Habbena and Stahl, 2011). Cette

caractéristique commune aux espèces *N. gargensis* (Hatzenpichler *et al.*, 2008) et *N. devanattera* (Lehtovirta-Morley *et al.*, 2011) est compatible avec le statut oligotrophe des milieux marins, habitats où les proportions d'AOA sont les plus importantes et les concentrations en ammonium plus faibles (Martens-Habbena *et al.*, 2009). Dans ce type d'écosystèmes, les AOA seraient fortement compétitives pour utiliser l'oxydation de l'ammonium comme source d'énergie.

Les AOA retrouvées dans les sols, telles que *N. viennensis* ou *N. koreensis*, sont elles, adaptées à des concentrations plus élevées en substrat (Jung *et al.*, 2011; Tourna *et al.*, 2011). Cependant, leur croissance est inhibée à des concentrations de 10 à 20 mM d'ammonium soit des seuils plus faibles que ceux tolérés par les AOB (50 à 1000 mM).

2.5.2 Importance de la salinité et du pH

La salinité apparaît dans de nombreuses études comme le facteur déterminant dans la régulation des communautés oxydant l'ammonium. Une étude conduite sur des lacs tibétains de haute altitude et de salinité variable a suggéré pour la première fois une différenciation phylogénétique des AOA selon un gradient de salinité (Hu *et al.*, 2010). D'autres travaux montrent que dans des écosystèmes estuariens, une salinité croissante induit une augmentation du nombre d'AOA (Mosier and Francis, 2008; Santoro *et al.*, 2008). Pourtant, il semblerait qu'il ne s'agisse pas du seul paramètre intervenant dans ce phénomène dans la mesure où de nombreux autres facteurs, tels que le pH, pourraient également être impliqués (Moin *et al.*, 2009). Un pH acide ($\text{pH} < 5$) pourrait favoriser une abondance et une activité plus importante des AOA comparées aux AOB (Gubry-Rangin *et al.*, 2011; Lu *et al.*, 2012; Nicol *et al.*, 2008). Ceci s'explique par le fait qu'une diminution du pH diminue la disponibilité en ammonium (De Boer *et al.*, 1991), bien qu'elle augmente dans le même temps la toxicité de l'acide nitreux (HNO_2). Par conséquent, comme les AOA ont généralement une plus forte affinité pour l'ammonium, elles sont favorisées dans ce type de milieux. C'est également ce qui pourrait expliquer que les représentants cultivés d'AOB ont une activité nulle ou limitée pour des pH inférieurs à 6 (Gubry-Rangin *et al.*, 2010).

En conclusion, l'établissement et la distribution des AOA dans les écosystèmes sont conditionnés par la combinaison de plusieurs facteurs, déterminant ainsi quelle sera la niche écologique la plus propice au développement de ces microorganismes.

2.6 Conclusions et hypothèses

Comme le démontre l'analyse bibliographique présentée précédemment, les *Archaea* sont des acteurs essentiels des écosystèmes tant en anoxie que dans les zones oxygénées. Longtemps ignorées dans les zones euphotiques des écosystèmes aquatiques, elles se révèlent être des acteurs clés de cycles biogéochimiques essentiels tels que celui de l'azote et du carbone. Cette prise de conscience a été rendue possible essentiellement grâce à l'essor considérable des techniques de biologie moléculaire durant ces deux dernières décennies. En effet, il y a encore quelques années, à l'initiation de ce projet de thèse, les données concernant leur diversité, leur structure, leur activité ou encore les facteurs les régulant étaient relativement limitées. Aujourd'hui encore, malgré une augmentation significative du nombre d'articles les concernant, de nombreuses contradictions et zones d'ombre persistent. Or, une meilleure connaissance de ces microorganismes, et plus particulièrement de ceux occupant des fonctions clés comme la nitrification, est aujourd'hui indispensable pour mieux cerner l'étendue réelle de leur rôle et de leur importance dans le fonctionnement des écosystèmes.

Une des découvertes les plus inattendues réalisées ces dernières années concerne la mise en évidence de l'implication des *Thaumarchaeota* dans une étape clé du cycle de l'azote, la nitrification. Cependant, bien que l'implication des *Archaea* (AOA) dans ce processus soit aujourd'hui acceptée, beaucoup d'incertitudes subsistent sur leur structuration, leur potentiel métabolique réel et les voies empruntées. Par ailleurs, les données actuelles semblent indiquer que les *Archaea* seraient très diversifiées et constituées d'écotypes qui se seraient adaptés aux conditions contrastées rencontrées dans les différents écosystèmes aquatiques, tels que les lacs, les estuaires ou encore les milieux marins. Ces populations pourraient également être dotées de capacités métaboliques encore insoupçonnées et différentes de celles connues aujourd'hui.

Pour répondre au moins partiellement aux questions toujours en suspens, de nombreux outils de génomique et de post-génomique ont été conçus ou améliorés ces dernières années. Ils permettent désormais de répondre à des problématiques d'écologie microbienne portant sur des environnements complexes. Parmi ces outils, certains dits à « haut débit » offrent la possibilité d'avoir une vision plus intégrative de l'ensemble des événements se déroulant dans un écosystème grâce à des approches dites de métagénomique ou de métagénétique.

Au moins en partie grâce à l'utilisation de ces outils, plusieurs hypothèses émanant des données bibliographiques ont pu être explorées lors de ce travail de thèse à savoir :

- l'hypothèse selon laquelle les *Archaea* (qu'il s'agisse de *Thaumarchaeota* ou d'*Euryarchaeota*) pourraient être constituées d'écotypes dont la distribution spatiale au sein d'un même écosystème varierait en fonction de la saison ;
- l'hypothèse selon laquelle la distribution et l'activité de ces écotypes d'*Archaea* seraient liées à un trait métabolique particulier. Dans le cas particulier des *Thaumarchaeota*, il pourrait s'agir du phénomène d'oxydation de l'ammonium, dont la régulation serait dépendante de facteurs physico-chimiques, tels que la salinité ou la concentration en ammonium, mais aussi de la compétition avec d'autres microorganismes planctoniques tels que les AOB. Ainsi, l'importance relative des AOA par rapport aux AOB serait susceptible de varier là aussi dans le temps et dans l'espace, mais également en fonction des écosystèmes.

SITES D'ETUDE ET APPROCHES METHODOLOGIQUES

3.1 Description des sites d'étude

Parmi les sites d'étude échantillonnés, plusieurs font l'objet d'un suivi régulier grâce à des observatoires. C'est le cas des lacs Pavin et Bourget suivis dans le cadre du programme SOERE GLACPE (Système Organisé en Réseau, Grands Lacs Péralpins) et du milieu côtier suivi dans le cadre du réseau SOMLIT (Service d'Observation en Milieu Littoral). L'objectif de tels observatoires est d'observer, de comprendre, et de modéliser l'évolution de l'état et des fonctionnements écologiques d'écosystèmes aquatiques soumis à un changement des pressions d'anthropisation locales et aux changements globaux en cours. Il s'agit d'outils essentiels pour comprendre le fonctionnement et l'évolution des écosystèmes en réponse à des changements environnementaux car ils permettent des surveillances sur des temps plus ou moins longs. Il existe à l'heure actuelle une émergence d'observatoires microbiologiques dans la communauté internationale qui ont pour objectifs (i) d'explorer la biodiversité des microorganismes, d'appréhender les facteurs qui contrôlent leurs changements et (ii) d'utiliser la veille microbiologique comme indicatrice de perturbations environnementales à court ou long terme.

3.1.1 Les milieux d'eau douce

Quatre milieux lacustres ont été échantillonnés lors de ce travail, trois localisés dans la région Auvergne (i.e. le lac d'Aydat, le lac Pavin et le réservoir de la Sep), le dernier étant localisé dans la région Rhône-Alpes (i.e. le lac Bourget). Le choix s'est porté sur ces milieux en raison de leurs particularités écologiques à savoir des concentrations en nutriments variables engendrant un statut trophique différent, des profondeurs et des dynamiques particulières mais aussi des pressions anthropiques variées.

3.1.1.1 Le lac d'Aydat

Le lac d'Aydat est situé à 25 km au sud-ouest de Clermont-Ferrand dans la partie méridionale de la chaîne des Puys, à 825 m d'altitude (Figure 12). Il s'est formé il y a près de 7500 ans suite au barrage de la Veyre (cours d'eau qui l'alimente désormais dans sa partie sud-ouest) par une coulée basaltique due à l'éruption des volcans jumeaux de la Vache et de Lassolas (Ogier, 1999). Ce lac a une profondeur maximale de 14.5 m et une superficie de 6.03

km². Il possède un bassin versant d'une superficie de 30.10^6 m². L'activité humaine autour du lac est essentiellement agricole avec une très nette prédominance de l'élevage en amont du lac. A cela s'ajoute le tourisme, relativement bien développé en période estivale suite à l'installation d'une base nautique. Le lac est typiquement eutrophe (Ogier, 1999). D'avril à octobre, la couche d'eau correspondant à l'hypolimnion est anoxique (Ogier, 1999).

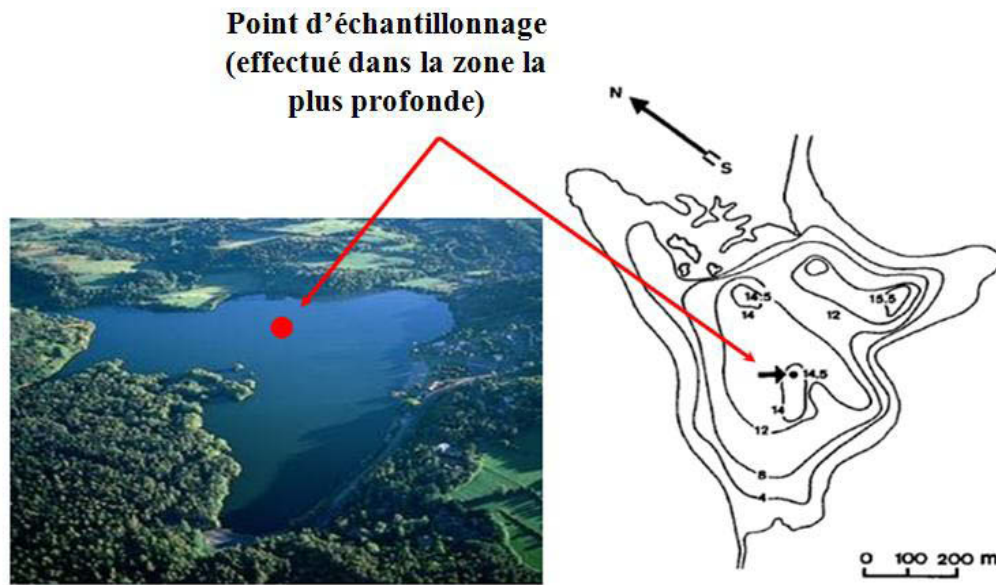


Figure 12 : Photographie illustrant le lac d'Aydat, et carte bathymétrique associée (http://domenicus.malleotus.free.fr/m/lac_aydat.htm).

3.1.1.2 Le lac Pavin

Le lac Pavin, comme le lac d'Aydat, est un lac de moyenne montagne. Distant d'environ 30 km de ce dernier, il est vieux d'environ 6500 ans. Il est situé sur le flanc Nord-Est du Puy de Montchal (Massif des Monts Dore) à une altitude de 1197 m (Figure 13). C'est un lac de cratère de profondeur maximale de 92 m et d'une superficie de 44 ha. Il s'est formé suite à une explosion violente produite par la rencontre du magma ascendant et d'une nappe d'eau (Maar ou nappe phréato-magmatique). Il est enchâssé dans des coulées de pyroclastites préexistantes et est alimenté par différentes sources superficielles (Fontaine du Loup, Source des prêtres...) auxquelles viennent s'ajouter l'apport de sources sous-lacustres. C'est un lac méromictique. Sa méromicticité est liée à l'importance de sa profondeur par rapport à sa surface. L'absence de brassage de la zone profonde conduit à la présence d'une zone anoxique permanente, le monimolimnion. Le lac Pavin est un lac oligomésotrophe (Boucher *et al.*, 2006) qui fait l'objet depuis de nombreuses années de nombreuses études conduites tant par

des microbiologistes, que des chimistes ou biogéochimistes (Borrel *et al.*, 2012; Chapron *et al.*, 2010; Colombet *et al.*, 2009).

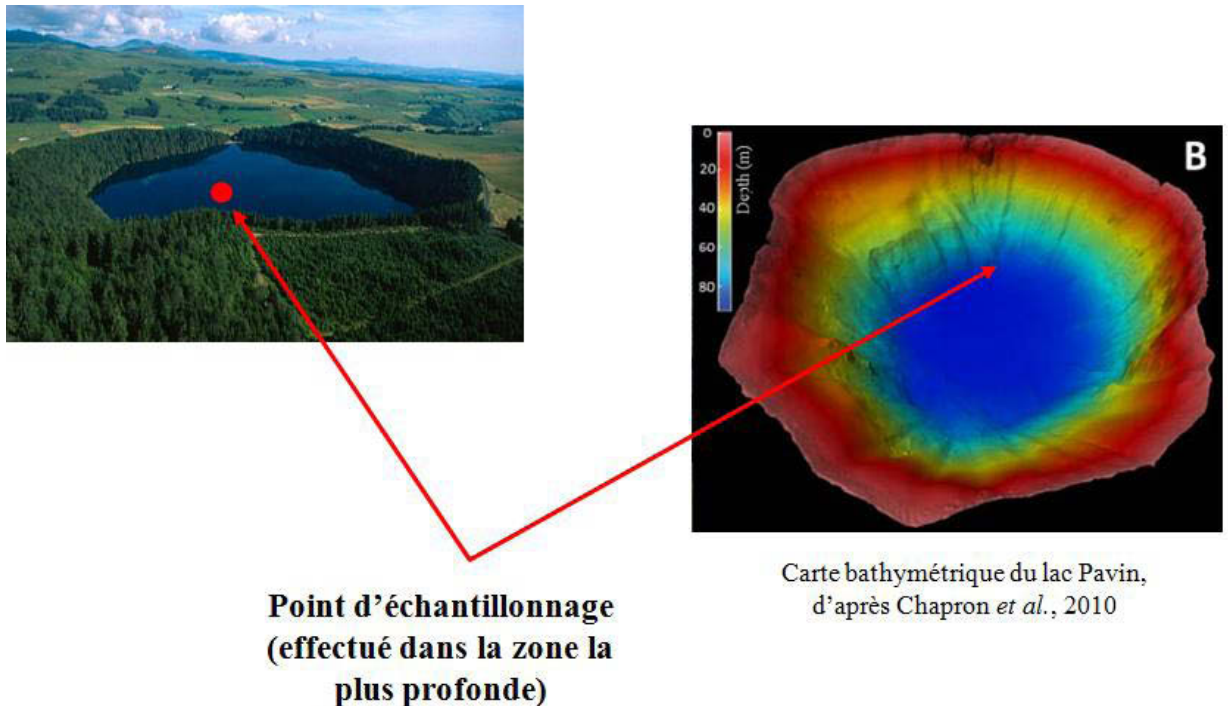


Figure 13 : Photographie illustrant le lac Pavin, et carte bathymétrique associée (<http://hydrologie.org/glu/ZZ/GP0782.HTM>).

3.1.1.3 Le réservoir de la Sep

Contrairement aux deux lacs précédents, le réservoir de la Sep est un lac artificiel construit en 1994 à des fins d'irrigation des terres de la Haute-Morge, espaces dédiés à l'agriculture (Figure 14). Le réservoir est situé à 500 m d'altitude et présente une profondeur maximale de 37 m pour une superficie de 0,33 km². Le bassin versant occupe une superficie de 27 km² et le temps moyen de résidence des eaux est de 220 jours. Le réservoir est vidangé régulièrement à des fins d'irrigation, surtout en été pour éviter la désoxygénation de la couche profonde. Cette retenue a été classée comme oligomésotrophe (Lepère, 2007).

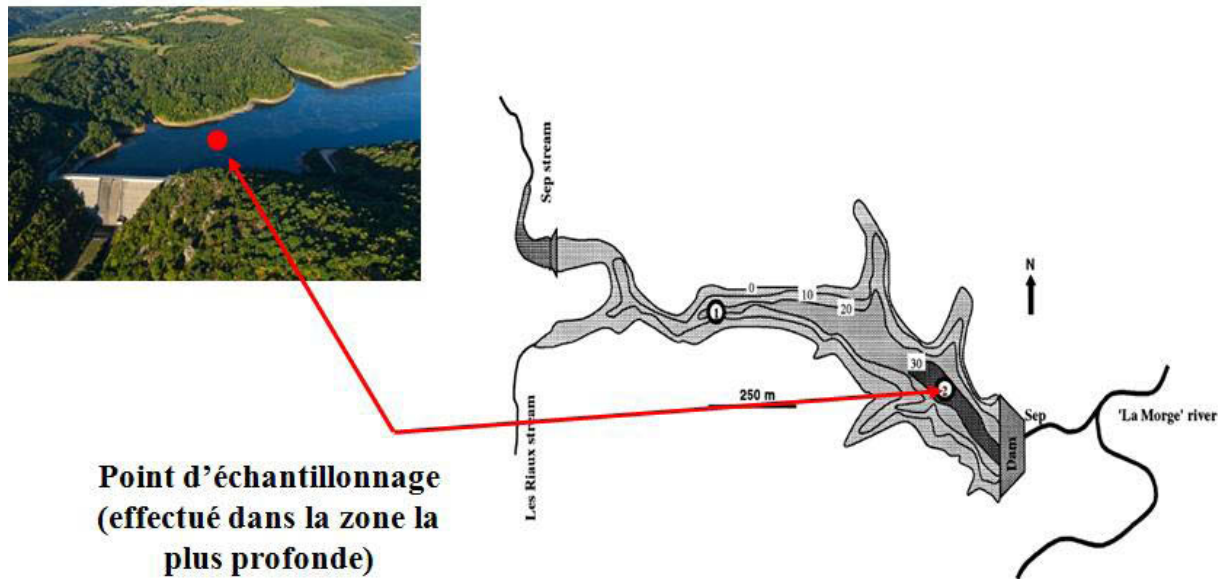


Figure 14 : Photographie illustrant le réservoir de la Sep, et carte bathymétrique associée (http://www.photo-paramoteur.com/photographies-aeriennes/auvergne-puy-de-dome/content/barrage-de-la-sep-22_large.html).

3.1.1.4 Le lac du Bourget

Situé dans le massif du Jura, le lac du Bourget (Figure 15) est situé au pied des Alpes à une altitude de 231.5 m. Il est considéré comme le plus grand lac naturel français (44.5 km² et 3.6 milliards de m³ d'eau). Formé après la dernière glaciation du Würm il y a environ 19000 ans, ce lac présente une profondeur maximale de 145.4 m (profondeur moyenne de 81 m). La nature essentiellement calcaire du substrat géologique du bassin versant conditionne la composition des eaux, bicarbonatées et calciques. Il est alimenté par trois rivières, la Lysse (68% des apports), le Sierroz (14% des apports) et le Tillet (18% des apports), et se déverse dans le Rhône au nord, par le canal de Savières. Il est délimité à l'ouest par les derniers contreforts du Jura méridional (chaîne de l'Epine, point culminant : Dent du Chat, 1390 m) et à l'est par le massif des Bauges, les montagnes de Cessens, de la Chambote, de Corsuet, et la colline de Tresserve. La durée de renouvellement des eaux du lac est de l'ordre de 7 ans. Le bassin versant du lac du Bourget couvre une superficie totale de 560 km² et ses altitudes maximale et moyenne sont respectivement de 1845 et 700 m. Ce lac a subi un processus d'eutrophisation jusqu'à la mise en place d'un important programme de restauration dans les années 1980. Il est actuellement considéré comme étant un lac mésotrophe (Lepère, 2007).

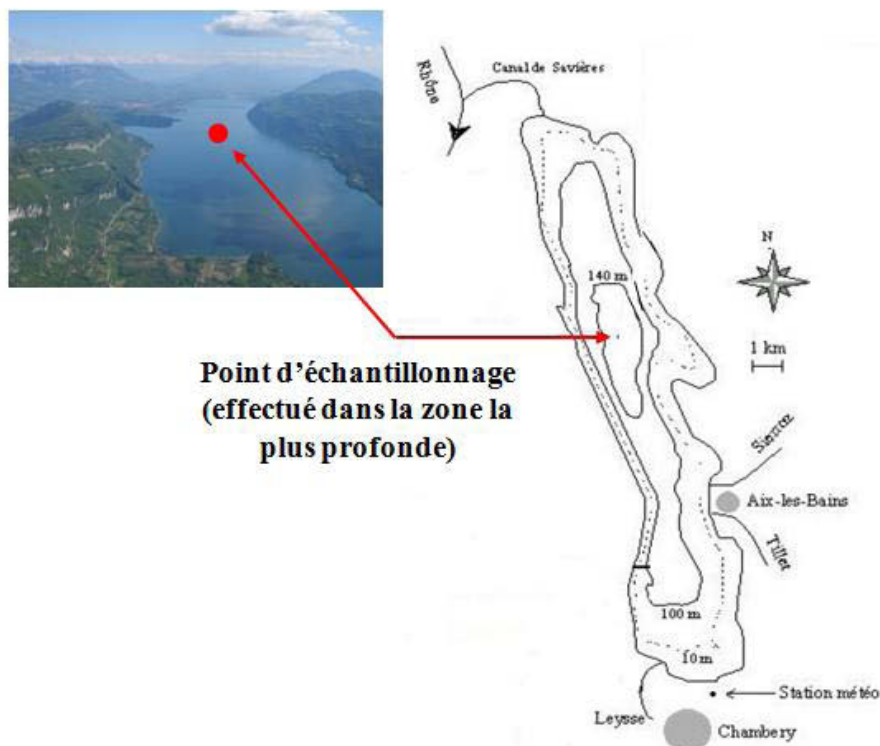


Figure 15 : Photographie illustrant le lac du Bourget, et carte bathymétrique associée (<http://www.lacdubourget.eu/plan-access-lac-du-bourget.html>).

3.1.2 Estuaire et milieu côtier

Parallèlement aux milieux d'eau douce, des milieux présentant des niveaux de salinité différents ont également été considérés lors de ce travail. Il s'agissait d'un milieu côtier situé en baie de Banyuls sur Mer et de l'estuaire de la Charente.

3.1.2 La baie de Banyuls sur Mer

Située en mer Méditerranée, la baie de Banyuls est équipée d'une station d'observation, celle du laboratoire Arago (station SOLA). Elle a été mise en place au début de l'année 1997 et fixée par 27 m de profondeur (Figure 16), en association avec l'initiative SOMLIT (Service d'Observation en Milieu LITtoral – <http://somlit.epoc.u-bordeaux1.fr/fr/>). Ce point est sous l'influence des apports fluviaux en provenance du Rhône, des fleuves côtiers tels que le Tech ou la Têt et localement des crues épisodiques de la Bailleurie en baie de Banyuls sur Mer.

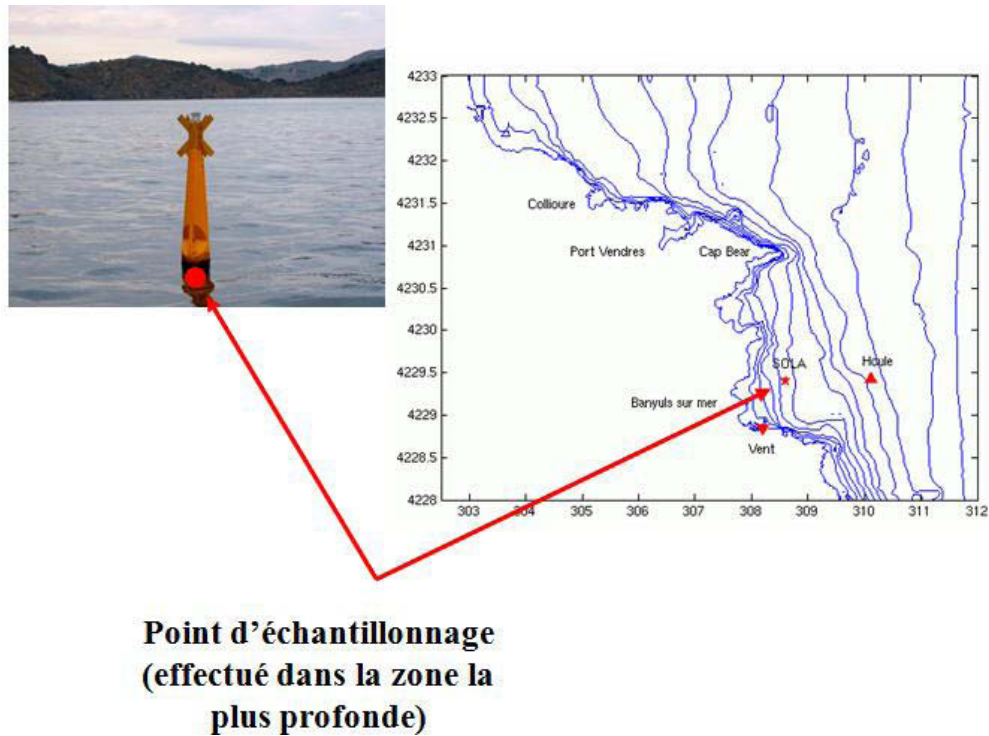


Figure 16: Photographie illustrant la station d'échantillonnage SOLA, et carte bathymétrique associée (<http://somlit.epoc.u-bordeaux1.fr/fr/>).

3.1.3 L'estuaire de la Charente

Le fleuve Charente prend sa source aux confins des départements de la Haute-Vienne et de la Charente à une altitude avoisinant 300 m. Il parcourt 360 km, de sa source jusqu'à l'océan Atlantique où il se jette au niveau de Fouras dans le bassin de Marennes Oléron (Figure 17). Il draine un bassin versant de 10000 km² et son estuaire est large de 3 km à son embouchure. La très faible pente dans la partie inférieure a induit la formation de nombreux méandres. Son bassin versant, à dominante rurale, présente une agriculture très diversifiée.

Plusieurs stations d'échantillonnage ont été positionnées le long de l'estuaire, caractérisées par des niveaux de salinité différents. Ainsi, le point d'échantillonnage de St-Savinien présente une salinité nulle alors que le dernier point de prélèvement situé à l'embouchure possède une salinité proche de celle retrouvée en mer (Figure 17).

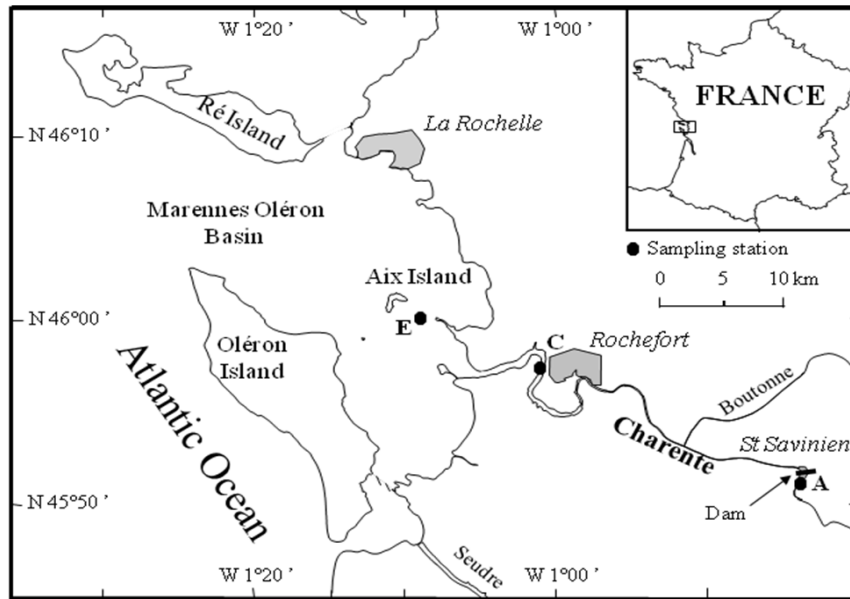


Figure 17 : Carte de l'estuaire de la Charente et des différents points de prélèvements échantillonnés. La station A est qualifiée de station d'eau douce, avec une salinité nulle, tandis que la station C dite mésohaline a une salinité moyenne de 14.9 PSU et la station E dite marine a une salinité moyenne de 33.2 PSU, d'après Auguet *et al.* (2005).

3.2 Approches d'écologie moléculaire

En écologie microbienne, le pyroséquençage et l'Illumina/Solexa sont des techniques très utilisées, ciblant ou non un gène d'intérêt. Néanmoins, les techniques NGS ne sont pas considérées comme étant strictement quantitatives et pour déterminer l'abondance réelle et la dynamique de groupes microbiens d'intérêt, il est nécessaire d'avoir recours à des techniques dédiées. C'est ce qui a été fait lors ce travail grâce aux approches de PCR en temps réel (qPCR et RT-qPCR) pour quantifier l'abondance des *Archaea*, des AOA et des AOB présents dans les milieux étudiés soit sur la base de marqueurs phylogénétiques (ADNr 16S) soit sur celle de marqueurs fonctionnels (*amoA*).

3.2.1 Conservation des échantillons et extraction des acides nucléiques

Les échantillons ont été prélevés dans la colonne d'eau à l'aide d'une bouteille Niskin. Pour préserver l'intégrité des ARN, une solution de RNA Later (sulfate d'ammonium 7,93 M, citrate de sodium 0,025 M, EDTA 0,02 M dans un volume final de 1,5 L d'eau RNase free, le

tout à pH 5,2, Lami *et al.* (communication personnelle)) a été ajoutée à l'eau échantillonnée (V/V).

Pour chaque prélèvement, les microorganismes sont concentrés par filtration avant de procéder à l'extraction des acides nucléiques. Brièvement, une pré-filtration est d'abord réalisée sur des filtres en polycarbonate de 5 µm de porosité (TMTP, Millipore) afin d'éliminer la fraction d'organismes supérieure à cette taille. Les microorganismes contenus dans le filtrat sont ensuite récoltés par une nouvelle filtration sur des filtres en polycarbonate de 0,2 µm de porosité (GTTP, Millipore) qui sont ensuite stockés à -80°C jusqu'à utilisation. Pour les milieux marins, les volumes traités étant beaucoup plus importants, les extractions ont été réalisées sur des cartouches Sterivex, à partir desquelles les acides nucléiques ont ensuite été extraits.

Brièvement, les microorganismes présents sur filtres ou cartouches Sterivex subissent dans un premier temps une lyse mécanique (cycles de congélation à -80°C/décongélation) puis une lyse enzymatique via l'utilisation d'un mélange de lysozyme, protéinase K, SDS et β-mercaptoéthanol. Les ADN et ARN sont alors co-extraits à l'aide du kit AllPrep DNA/RNA de Qiagen selon les instructions du fournisseur. L'ADN génomique (ADNg) est dosé par spectrofluorimétrie (ND 1000 ; Nanodrop). Les ARN totaux, sont testés individuellement par PCR pour évaluer une potentielle contamination par de l'ADNg avant d'être rétrotranscrits. La technique de rétrotranscription a été adaptée aux différents gènes ciblés. Ainsi, pour l'étude du marqueur ADNr 16S, les ARN ont été rétrotranscrits *via* l'utilisation d'un mélange d'hexamères générés de façon aléatoire et du kit Superscript VILO d'Invitrogen, selon les préconisations du fournisseur. Pour les gènes fonctionnels, la rétrotranscription des ARNm a été initiée à l'aide d'une amorce ciblant spécifiquement le gène d'intérêt (amorce R ou reverse) et du kit Superscript III d'Invitrogen.

Les ADNg des échantillons lacustres du lac d'Aydat, du réservoir de la Sep et du lac du Bourget utilisés dans le Chapitre 2 (Article 3), ont été extraits par la technique phénol/chloroforme/alcool isoamylique.

3.2.2 Amplification par PCR des gènes cibles

Les gènes d'ADNr 16S et fonctionnels étudiés ont été amplifiés par PCR à l'aide de couples d'amorces choisies en fonction de l'application souhaitée, à savoir le séquençage

Sanger, le pyroséquençage et la PCR en temps réel (qPCR et qRT-PCR). Ils sont, pour la grande majorité, issus de la bibliographie (Table 2).

Table 2. Liste des gènes étudiés et des amorces utilisées en PCR pour des fins de séquençage Sanger, pyroséquençage et PCR en temps réel.

Application	Gènes	Amorce	Séquence (5' – 3')	Température d'hybridation	Référence	Chapitre d'application
PCR / Sanger	ADNr 16S d' <i>Archaea</i>	Ar4f	TCYGGTTGATCCTGCCRG	56°C	Jurgens <i>et al.</i> , 2000	Chapitre 2, Article 3
		Ar958r	YCCGGCGTTGAVTCCAATT		Jurgens 2002	
	<i>amoA</i> de <i>Thaumarchaeota</i>	Arch-amoAF	STAATGGTCTGGCTTAGACG	53°C	Francis <i>et al.</i> , 2005	
		Arch-amoAR	GCGGCCATCCATCTGTATGT			
	<i>amoA</i> de β-protéobactéries	amoA-1F	GGGGTTTCTACTGGTGGT	56°C	Rotthauwe <i>et al.</i> , 1997	
		AmoA-RNEW	CCCCTCBGSAAVCCTTCTTC			
PCR / Pyroséquençage	ADNr 16S d' <i>Archaea</i>	Arch349F	GYGCASCAGKCGMGAAW	59°C	Takai <i>et al.</i> , 2000	Chapitre 1 et 2, (Articles 1 et 4)
		Arch 806R	GGACTACVSGGTATCTAAT	57°C	Herfort <i>et al.</i> , 2009 Casamayor <i>et al.</i> , 2002	Chapitre 1, Article 2
		Arch519F	CAGCCGCCGCGGTAA			
		Arch915R	GTGCTCCCCGCCAATTCCT			
	<i>amoA</i> de <i>Thaumarchaeota</i>	Arch-amoAF	STAATGGTCTGGCTTAGACG	53°C	Francis <i>et al.</i> , 2005	Chapitre 2, Article 4
		Arch-amoAR	GCGGCCATCCATCTGTATGT			
	<i>amoA</i> de β-protéobactéries	amoA-1F	GGGGTTTCTACTGGTGGT	56°C	Rotthauwe <i>et al.</i> , 1997	Chapitre 2, Article 4
		Bacamo2R	CCCCTCKGSAAAGCCTTCTTC			
qPCR / q-RT-PCR	ADNr 16S de <i>Thaumarchaeota</i>	Lake 386F	ACGACWRGGTYARTCCGAGTGRTT	60°C	Hugoni <i>et al.</i> 2013	Chapitre 2, Article 3
		957R	CGGCGTTGACTCCAATTG		Ochsenreiter <i>et al.</i> , 2003	
	<i>amoA</i> de <i>Thaumarchaeota</i>	Arch-amoA-for	CTGAYTGGGCTGGACATC	58,5°C	Wuchter <i>et al.</i> , 2006	Chapitre 2, Article 3
		Arch-amoA-rev	TTCTTCTTTGTTGCCAGTA			
		CrenAmoAModF	TGGCTAAGACGMTGTA	52°C	Mincer <i>et al.</i> , 2007	Chapitre 1 et 2, (Articles 2 et 4)
			CrenAmoAModR			
	<i>amoA</i> de β-protéobactéries	amoA-1F	GGGGTTTCTACTGGTGGT	56°C	Rotthauwe <i>et al.</i> , 1997	Chapitre 1 et 2, (Articles 2 et 4)
		AmoA-RNEW	CCCCTCBGSAAVCCTTCTTC			
	<i>ureC</i> de <i>Thaumarchaeota</i>	Thaum-UreC F	ATGCAATYTGTAATGGAACWACWAC	55°C	Alonso-Saez <i>et al.</i> , 2012	Chapitre 2, Article 4
		Thaum-UreC R	AGTTGTGCCCAATCTTCATGTAATTTTA			

Nos choix, pour ce qui concerne l'approche clonage-séquençage Sanger se sont portés sur des amorces couramment utilisées. Dans le cadre du pyroséquençage, le choix des amorces ciblant les ADNr 16S d'*Archaea* a été conditionné par des tests d'amplifications qui ne nous ont pas permis d'utiliser le même couple pour tous les écosystèmes étudiés (ces résultats seront discutés dans la partie Discussion 6.1.1). La taille des amplicons est voisine de 400 pb (369 pb pour le couple 915R-519F et 457 pb pour le couple 806R-349F). Pour le

pyroséquençage des amplicons de gènes *amoA*, les choix méthodologiques du prestataire externe ont guidé nos travaux. Enfin, en ce qui concerne la PCR en temps réel, la taille optimale des fragments amplifiés doit théoriquement être comprise entre 150 et 400 pb. Néanmoins, l'amplification des gènes d'ADNr 16S de *Thaumarchaeota* a été réalisée grâce à une amorce disponible dans la littérature (Ochsenreiter *et al.*, 2003) et la seconde a été dessinée lors de l'étude décrite dans le Chapitre 2 (Article 3), générant un fragment de 571 pb. Lors de cette même étude, nous avons quantifié les gènes *amoA* archéens à l'aide des amorces de Wuchter *et al.* (2006), mais suite aux critiques de ces amorces relevées dans la littérature (mésappariements et spécificité mis en cause (Church *et al.*, 2009)), nous avons choisi, dans les Chapitres 1 (Article 2) et 2 (Article 4), d'utiliser le couple préconisé par Mincer *et al.* (2007), qui présentait de surcroît une taille plus adaptée à la qPCR (150 pb).

3.2.2.1 La PCR quantitative

La PCR quantitative est une technique permettant de quantifier rapidement un gène d'intérêt dans de nombreux échantillons. Cette technique est également applicable à la quantification des transcrits d'un gène cible dans un échantillon, mais nécessite alors une étape préalable de rétrotranscription (RT) permettant de travailler sur les ADN complémentaires (ADNc).

Cette approche repose sur une amplification PCR classique couplée à la quantification d'un signal fluorescent. L'intensité de fluorescence à chaque cycle est directement proportionnelle à la quantité d'amplicons produite, elle-même directement proportionnelle à la quantité initiale d'ADN présente dans le milieu réactionnel. Lors de cette thèse, nous avons choisi d'opter pour une quantification dite absolue, qui se base sur l'utilisation d'une gamme étalon construite à partir de plasmides contenant le gène d'intérêt à des concentrations initiales connues. Pour les différents gènes ciblés au cours de cette thèse, les gammes de calibration ont été établies à partir d'un mélange de plusieurs plasmides contenant un insert du gène ciblé et non à partir d'un seul afin de limiter les biais engendrés par l'étude d'échantillons environnementaux.

La méthode au SYBR Green (Figure 18) est une des approches les plus utilisées et celle sur laquelle s'est porté notre choix. Le SYBR Green est un intercalant qui va venir se fixer dans la molécule d'ADN et va émettre une fluorescence. Par conséquent, l'intensité de fluorescence sera proportionnelle à la quantité d'ADN présente générée (Figure 18).

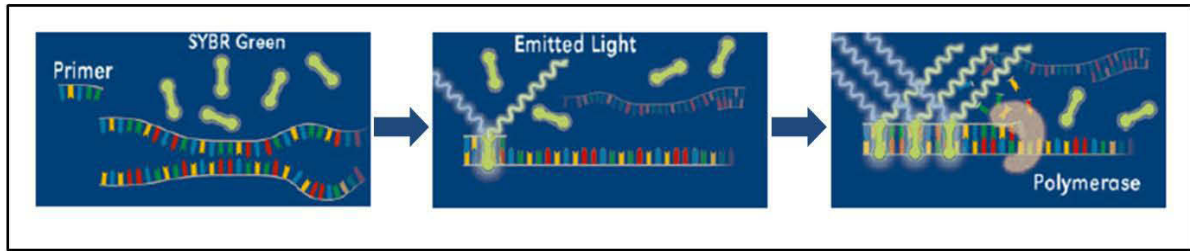


Figure 18 : Réaction d'incorporation du SYBR Green durant la réaction de PCR (d'après le site : <http://homepages.strath.ac.uk/~dfs97113/BB310/lect403.html>)

Les gènes ciblés au cours de cette étude étaient les gènes d'ADNr 16S archéens et *amoA* de bactéries et d'*Archaea* présents dans des échantillons d'ADNg et d'ADNc environnementaux.

3.2.2.2 Le pyroséquençage d'amplicons ou métagénétique

La technologie de séquençage 454 autorise l'analyse simultanée de plusieurs échantillons par l'ajout d'un code barre ou tag lors de l'étape d'amplification des ADN génomiques par PCR. Les amorces utilisées sont dites fusionnées (Figure 19) car elles sont le fruit d'une construction incluant un adaptateur (le même pour tous les échantillons) nécessaire au pyroséquençage, un tag (propre à chaque échantillon) et l'amorce sens (ou forward). Ainsi, autant d'amplicons que d'étiquettes disponibles peuvent être préparés, ils seront par la suite dosés et mélangés de façon équimolaire avant pyroséquençage, ceci afin de séquencer équitablement chaque amplicon. Dans le cadre de nos travaux, nous avons choisi un séquençage unidirectionnel, à partir de l'adaptateur A (Figure 19).

En outre, en ce qui concerne les études conduites en milieu côtier (Chapitre 1, Article 1) et estuarien (Chapitre 2, Article 4), la prestation a été réalisée dès l'étape de PCR par le prestataire externe (Research and Testing Laboratory, Lubock et MR. DNA, Shallowater, Texas, Etats Unis), tandis que dans le cadre de l'étude des milieux lacustres (Chapitre 1, Article 2), nous avons choisi de préparer les amplicons. Pour ce faire, les amorces fusionnées ont été synthétisées en utilisant les tags préconisés par Roche (http://www.liv.ac.uk/media/livacuk/centreforgenomicresearch/The_GS_FLX_Titanium_Chemistry_Extended_MID_Set.pdf).

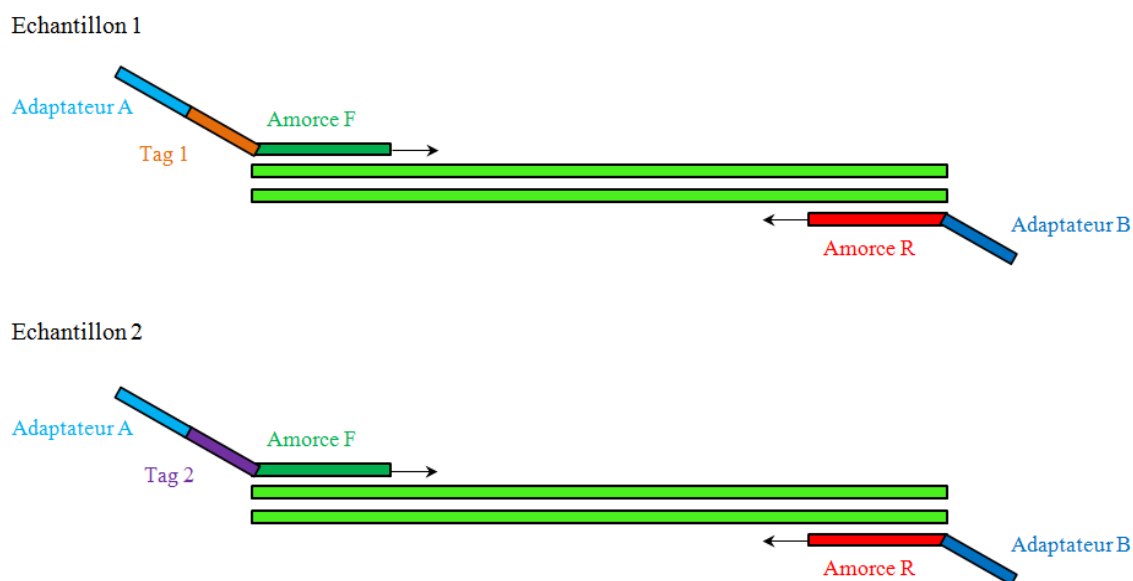


Figure 19 : Schéma représentant les constructions des amorces utilisées pour le pyroséquençage d'amplicons, d'après http://www.igsb.org/uploads/pdf/TCB-09013_AmpliconFusionPrimer_DesignGuidelines.pdf.

Par la suite, une amplification clonale des molécules d'ADN simple brin est réalisée après fixation de ces dernières à des microbilles possédant à leur surface l'amorce complémentaire à l'adaptateur A. Une mise en émulsion de cet ensemble (microbille-ADN simple brin) en présence des réactifs pour PCR est alors effectuée. Après amplification, les microgouttelettes sont dissociées et enrichies puis déposées sur une plaque en fibres optiques (plaque PicoTiter™) d'1,6 millions de puits. Les puits possèdent un diamètre qui assure le dépôt d'une microbille par puits. Dans chacun de ces puits est alors réalisé le séquençage des ADN simple brin selon le principe de pyroséquençage décrit ci-dessous. Brièvement, après incorporation d'un nucléotide dans le brin d'ADN en synthèse, une libération d'un pyrophosphate (PPi) se produit. Ce PPi est alors transformé en ATP par une ATP sulfurylase. Une luciférase couple alors cet ATP à une luciférine ce qui produit une oxyluciférine en émettant un signal lumineux capté par le scanner du séquenceur. Les nucléotides en surplus dans le milieu réactionnel sont alors dégradés par une apyrase.

Dans le cadre de ce travail de thèse, le pyroséquençage a été réalisé par des séquenceurs de deuxième génération par la technologie Roche 454 GS-FLX Titanium par différentes plateformes de séquençage développées en France (Plateforme GINA, Clermont Ferrand) et aux Etats-Unis (Research and Testing Laboratory, Lubock et MR. DNA Laboratory, Shallowater).

3.2.3 Traitement bioinformatique des données de NGS

Le traitement des données générées par les techniques de séquençages haut débit telles que le pyroséquençage nécessite le développement et l'utilisation d'outils bioinformatiques adaptés. Plusieurs outils destinés à ces traitements sont actuellement disponibles. C'est par exemple le cas de QIIME, qui permet entre autres de traiter les données de pyroséquençage d'amplicons. Cet outil assigne la taxonomie par Blast ou en utilisant le classifieur RDP, or il a été montré qu'un autre outil développé au sein de notre équipe, nommé PANAM (Phylogenetic Analysis of Next Generation Amplicons (Taib *et al.*, 2013)), permet de restituer une affiliation taxonomique plus fiable et de construire des phylogénies.

Nous avons donc choisi d'utiliser cet outil qui était initialement dédié à l'étude des amplicons de protistes (ADNr 18S). L'utilisation de ce logiciel a nécessité la construction préalable d'une base de données *Archaea*, nous permettant une affiliation taxonomique précise. Nous avons utilisé comme base de données de séquences de référence la base SSURef 108 de SILVA (Pruesse *et al.*, 2007; Quast *et al.*, 2013) pour laquelle les séquences d'*Archaea* retenues ont une bonne qualité d'alignement (>75%) et une taille supérieure ou égale à 900 pb. Nous avons également vérifié l'annotation de séquences faisant partie de clades définis en milieu d'eau douce ou marins. Les séquences sélectionnées, au nombre de 11028, sont ensuite séparées en 4 groupes monophylétiques correspondant aux *Crenarchaeota*, *Euryarchaeota*, *Thaumarchaeota* et « autres », constituant des profils taxonomiques nécessaires à l'affiliation taxonomique.

Une fois la base construite, le traitement commence par le nettoyage des séquences (Figure 20). Les tags et adaptateurs sont enlevés, et seules les séquences de plus de 200 pb, dont le score de qualité était supérieur à 25 ou 27 selon l'étude, l'amorce F retrouvée à 100%, et sans base « N », ont été retenues. Notre choix du score de qualité a été guidé par la volonté de conserver suffisamment de séquences dans les différents jeux de données (ADN et ARN) après normalisation, c'est pourquoi nous avons utilisé une valeur de 25 pour l'étude des milieux lacustres (Chapitre 1, Article 2) et du milieu estuarien (Chapitre 2, Article 4) qui étaient basées principalement sur une des analyses taxonomiques au niveau du clade. Pour l'analyse de la biosphère rare (Chapitre 1, Article 1) nous avons choisi d'utiliser une stringence plus forte (score de 27), permettant de conserver des séquences présentant le minimum d'ambiguïté quant à leur qualité. Dans le cadre de cette dernière étude (Chapitre 1, Article 1), la détection et l'élimination de séquences chimériques putatives a été réalisée en

utilisant Uchime (Edgar et al., 2011), et a permis de montrer qu'elles ne représentaient qu'environ 1,6% des séquences.

Les séquences nettoyées sont clusterisées en utilisant Uclust (Edgar, 2010). Le seuil de 97 % choisi est basé sur les données bibliographiques en relation avec la région ciblée pour le domaine des *Archaea* (Kim et al., 2011b). Les OTUs générées sont ensuite comparées avec la base de référence puis triées par grand groupe monophylétique en utilisant Usearch (Edgar, 2010). Enfin, ces alignements sont utilisés pour construire les phylogénies à l'aide de FastTree (Price et al., 2010) à partir desquelles la taxonomie de chaque OTU a été déterminée.

Toutes les analyses bioinformatiques ont été conduites en étroite collaboration avec Najwa Taib, doctorante dans l'équipe Microbiologie de l'Environnement du LMGE.

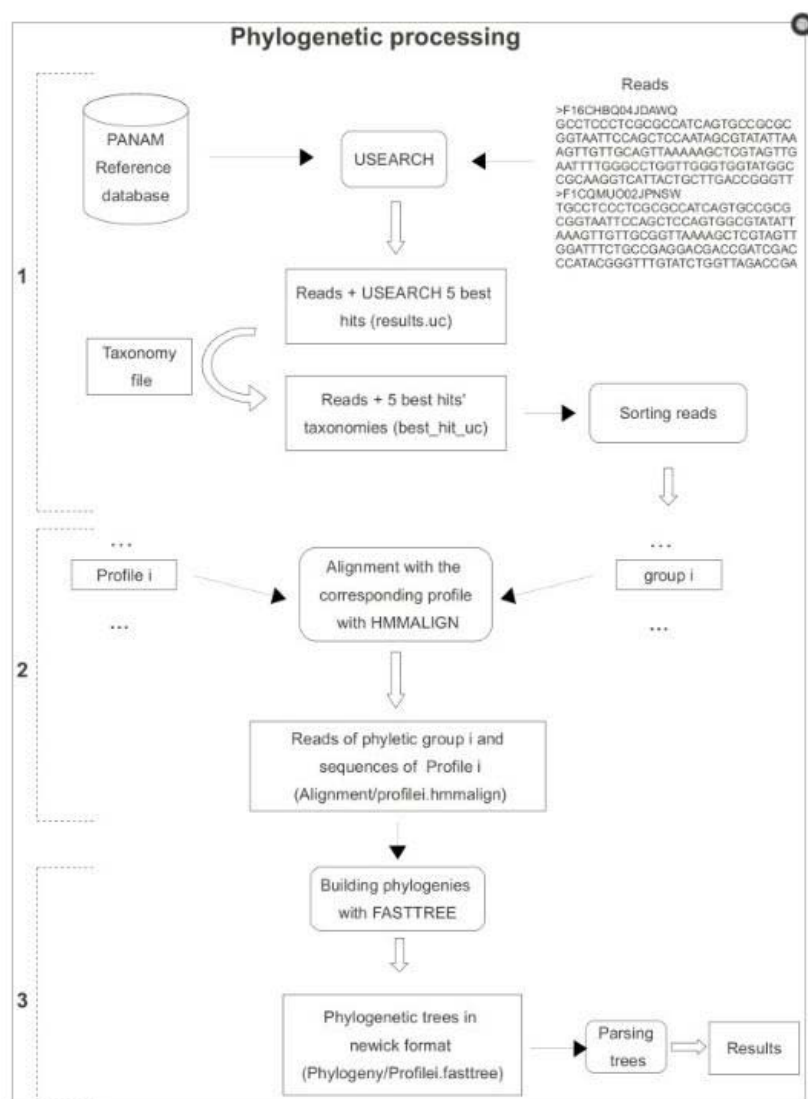


Figure 20 : Présentation des différentes étapes implémentées dans PANAM pour le nettoyage et l'affiliation taxonomique des séquences issues de NGS (Taib *et al.*, 2013).

CHAPITRE 1 : Structure des communautés d'*Archaea* actives dans les écosystèmes aquatiques

4.1 Contexte et objectifs

L'analyse bibliographique témoigne de l'ubiquité des *Archaea* dans la plupart des écosystèmes aquatiques oxiques. Dans ces milieux, qu'ils soient d'eau douce ou marins, le potentiel métabolique associé à ces microorganismes reste encore mal connu. A ce jour, l'implication des *Thaumarchaeota* dans le processus d'oxydation de l'ammonium est avérée ; les *Euryarchaeota* étant principalement associées au processus de méthanogenèse. Ces principales voies métaboliques bien décrites ne suffisent cependant pas à comprendre la dynamique des *Archaea* dans les écosystèmes. En effet, différentes études suggèrent la présence d'écotypes (groupes phylogénétiquement distincts ayant des rôles écologiques différents, (Beman *et al.*, 2008; Hallam *et al.*, 2006b; Mincer *et al.*, 2007)), au sein des principaux clades décrits (MGI, MEG, LDS...). L'ubiquité apparente de certains clades pourrait en fait cacher une microdiversité qui différencierait les populations d'*Archaea* dans le temps ou entre écosystèmes. D'autre part, les études conduites sur les communautés bactériennes témoignent de la richesse de la biosphère rare mais également de son activité dans les milieux aquatiques (Campbell *et al.*, 2011). La mise en évidence d'une biosphère rare chez les *Archaea* est un fait récent (Galand *et al.*, 2009) mais son activité et son rôle dans le fonctionnement des écosystèmes reste encore une hypothèse. Ainsi, les deux articles présentés dans ce chapitre ont pour objectif d'évaluer la structure et l'activité des communautés d'*Archaea* en lien avec des niches écologiques particulières liées aux dynamiques temporelles et spatiales des écosystèmes aquatiques.

Pour étudier les successions saisonnières, nous avons mené une étude à long terme (Article 1) correspondant à des observations échelonnées sur 3,5 ans en bénéficiant des données recueillies à l'observatoire océanologique de Banyuls sur Mer. Il est à noter que si que les milieux marins ont largement été étudiés (Agogué *et al.*, 2008; Church *et al.*, 2003; Francis *et al.*, 2005), aucune étude visant à évaluer la structure des communautés d'*Archaea* à long-terme n'avait été entreprise. Dans ce premier article, nous avons ainsi pu évaluer si la structuration des communautés actives était stable dans le temps, et si des récurrences interannuelles étaient observables. D'autre part nous avons pu tester, grâce à un jeu de données important tant au niveau qualitatif que quantitatif, différentes hypothèses liées à la biosphère rare.

Dans le second article, nous sommes focalisés sur les écosystèmes lacustres qui, contrairement aux milieux marins avaient fait l'objet de peu d'études à l'initiation de cette

thèse (Auguet *et al.*, 2011; Comeau *et al.*, 2012; Vissers *et al.*, 2013). Ce travail nous a permis d'évaluer si différentes niches écologiques pouvaient être mises en lien avec différentes communautés actives d'*Archaea*. Nous avons plus précisément caractérisé la structure des communautés actives d'*Archaea* dans deux zones favorables à leur activité : l'oxycline où les *Archaea* sont abondantes et connues pour réaliser l'oxydation de l'ammonium (Lliros *et al.*, 2010; Pouliot *et al.*, 2009) ; et la zone anoxique, lieu préférentiel pour la méthanogenèse.

Les approches mises en place pour caractériser les populations archéennes reposent essentiellement sur des outils « haut débit » et sur l'approche de métagénomique, à savoir le pyroséquençage d'amplicons appliqué à un marqueur phylogénétique : le gène codant pour l'ADNr 16S.

4.2 Article 1 : Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters

Mylène Hugoni & Najwa Taib, Didier Debroas, Isabelle Domaizon, Isabelle Jouan Dufournel, Gisèle Bronner, Ian Salter, Hélène Agogué, Isabelle Mary, and Pierre E. Galand
Proceedings of the National Academy of Sciences USA
2013

Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters

My lene Hugon^{a,b,1}, Najwa Taib^{a,b,1}, Didier Debroas^{a,b}, Isabelle Domaizon^c, Isabelle Jouan Dufournel^{a,b}, Gis le Bronner^{a,b}, Ian Salter^{d,e}, H l ne Agogue^f, Isabelle Mary^{a,b,2}, and Pierre E. Galand^{d,g}

^aLaboratoire "Microorganismes: G nome et Environnement," Clermont University, Universit  Blaise Pascal, F-63000 Clermont-Ferrand, France;

^bLaboratoire Microorganismes, G nome et Environnement, Centre National de la Recherche Scientifique (CNRS), Unit  Mixte de Recherche (UMR) 6023, F-63171 Aubi re, France; ^cInstitut National de la Recherche Agronomique, UMR 42 Centre Alpin de Recherche sur les R seaux Trophiques et  cosyst mes Limniques, F-74200 Thonon les Bains, France; ^dUniversit  Pierre et Marie Curie-Paris 6, UMR 8222, Laboratoire d' cog ochimie des Environnements Benthiques (LECOB), UMR 7621, Laboratoire d'Oc anographie Microbienne (LOMIC), Observatoire Oc anologique, F-66650 Banyuls-sur-Mer, France;

^eCNRS, UMR 7621, LOMIC, Observatoire Oc anologique, F-66650 Banyuls-sur-Mer, France; ^fLittoral, Environnement et Soci t s, UMR 7266, CNRS, University of La Rochelle, 17000 La Rochelle, France; and ^gCNRS, UMR 8222, LECOBS, Observatoire Oc anologique, F-66650 Banyuls-sur-Mer, France

Edited by David M. Karl, University of Hawaii, Honolulu, HI, and approved March 1, 2013 (received for review September 28, 2012)

Marine Archaea are important players among microbial plankton and significantly contribute to biogeochemical cycles, but details regarding their community structure and long-term seasonal activity and dynamics remain largely unexplored. In this study, we monitored the interannual archaeal community composition of abundant and rare biospheres in northwestern Mediterranean Sea surface waters by pyrosequencing 16S rDNA and rRNA. A detailed analysis of the rare biosphere structure showed that the rare archaeal community was composed of three distinct fractions. One contained the rare Archaea that became abundant at different times within the same ecosystem; these cells were typically not dormant, and we hypothesize that they represent a local seed bank that is specific and essential for ecosystem functioning through cycling seasonal environmental conditions. The second fraction contained cells that were uncommon in public databases and not active, consisting of aliens to the studied ecosystem and representing a nonlocal seed bank of potential colonizers. The third fraction contained Archaea that were always rare but actively growing; their affiliation and seasonal dynamics were similar to the abundant microbes and could not be considered a seed bank. We also showed that the major archaeal groups, Thaumarchaeota marine group I and Euryarchaeota group II.B in winter and Euryarchaeota group II.A in summer, contained different ecotypes with varying activities. Our findings suggest that archaeal diversity could be associated with distinct metabolisms or life strategies, and that the rare archaeal biosphere is composed of a complex assortment of organisms with distinct histories that affect their potential for growth.

ong-term dynamic | dormancy | taxonomic diversity | microbial observatory | Somlitt

The seasonal dynamics of marine microorganisms have traditionally been studied at the DNA level (1, 2), but recent studies have shown the importance of differentiating the active communities from the total communities (3–5). One method to explore an aspect of activity (i.e., the growth rate for specific taxa) is to investigate microbial communities with both 16S rRNA and 16S rDNA (6–8). The use of the 16S rRNA-to-rDNA sequence ratio as an index of microbial growth has revealed a generally positive correlation between abundance and activity in coastal surface bacterial communities (4, 9). However, abundant microbes are not always the most active (3), even though they contribute greatly to ecosystem functioning. An important finding is that growth can be detected among low-abundance taxa, also known as the rare biosphere (4, 7), which was first defined with the development of new sequencing technologies, allowing a deep coverage of the diversity of natural communities (10). Rare taxa have been hypothesized to consist of dormant microorganisms (or a seed bank) that could potentially be resuscitated under different environmental conditions (11). However,

the discoveries that the rare biosphere had a biogeography (12), and that a significant portion of the rare community was active (4, 7), with growth rates that decreased as abundance increased (4), suggest that the rare biosphere is not solely a dormant seed bank (13). A rare biosphere has been detected within the domain Archaea (12), and although we have begun to gain insights into the dominant archaeal phylotypes, the community structure of the rare Archaea remains largely uncharacterized.

Marine planktonic Archaea have been recently recognized as main drivers of the aerobic ammonia oxidation in many aquatic ecosystems, suggesting an important role in the nitrogen cycle (14–16). They have traditionally been described as spanning three major groups: Thaumarchaeota marine group (MG) I, which is more abundant in meso- and bathypelagic waters (17–19), Euryarchaeota MGII, which is more abundant in surface waters, and Euryarchaeota MGIII, which is restricted to deeper waters (20, 21). The diversity of Archaea is, however, much more complex; for instance, MGI appears to have distinct clusters segregated according to depth and location (22). A recent metagenomic characterization of MGI from north Atlantic coastal surface waters also suggested the presence of at least two dominant environmental populations that are divergent from each other (23). The presence of at least two clusters was also demonstrated in the Mediterranean Sea (24) and corresponded to groups previously detected in different oceanic provinces (20). Whether this taxonomic diversity corresponds to distinct ecotypes, i.e., groups of microorganisms playing distinct ecological roles and belonging to genetically cohesive and irreversibly separate evolutionary lineages (25), is not known because the relationship among archaeal activity, environmental conditions, and sequence abundance has never been studied. Moreover, the ecological control of archaeal diversity patterns over long time scales remains poorly understood (24).

By monitoring surface archaeal communities in monthly intervals during a 3.5-y period at the Banyuls-sur-Mer Bay Microbial Observatory, a site representative of the coastal northwest Mediterranean Sea, we aimed to describe the structure of the rare archaeal

Author contributions: I.D., H.A., I.M., and P.E.G. designed research; M.H. and N.T. performed research; M.H., N.T., D.D., I.S., and I.M. contributed new reagents/analytic tools; M.H., N.T., I.J.D., G.B., and P.E.G. analyzed data; M.H., N.T., I.M., and P.E.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The pyrosequencing data reported in this paper has been deposited in the Dryad database, <http://datadryad.org> (doi no. 10.5061/dryad.q5903).

¹M.H. and N.T. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: isabelle.mary@univ-bpclermont.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216863110/-DCSupplemental.

biosphere by testing whether it is composed of a seed bank of dormant cells or represents microorganisms with high growth rates. By targeting both 16S rRNA and rDNA, we also verified whether different archaeal clusters represent distinct ecotypes and assessed the seasonal activity dynamics of marine Archaea.

Results

Rare and Abundant Phylotypes. We used pyrosequencing to follow changes in the community structure, relative sequence abundance, and potential activity of Archaea over time. A total of 351 operational taxonomic units (OTUs) were retrieved in the 16S rDNA dataset, representing a total of 65,833 sequences. Seventeen OTUs were abundant (>1% and occurred in more than one sample) and contained 97% of all of the sequences. The 16S rRNA dataset was composed of 52,181 sequences consisting of 348 OTUs. Rarefaction curves for both 16S rDNA and 16S rRNA indicated that, in most cases, the sequencing depth captured the diversity present in the natural archaeal community (Fig. S1).

Only two OTUs were always abundant (OTU 2 affiliated with MGI, and OTU 13 affiliated with MGII.B), whereas the remaining 15 abundant OTUs were rare in some samples ($\leq 0.2\%$ of the sequences in a sample). Typically, OTUs abundant in winter became rare in summer and vice versa. All the abundant OTUs were active when they were abundant (Fig. 1A), and the plot of the 16S rDNA against 16S rRNA OTU frequencies had an intercept at zero and showed a high correlation between 16S rRNA and 16S rDNA (Kendall nonparametric $\tau = 0.7$; $P < 0.001$; $n = 224$). However, 16S rRNA and 16S rDNA were more poorly correlated when the abundant OTUs became rare ($\tau = 0.3$; $P < 0.001$; $n = 72$), with some OTUs showing high activity (16S rRNA/rDNA ratio >1) whereas others had low or no activity (16S rRNA/rDNA ratio <1; Fig. 1B).

All OTUs were compared with the entire SILVA database to ascertain if they were globally common (i.e., a high similarity to reference sequences) or uncommon (i.e., a low similarity). The abundant DNA OTUs were common, with an average 98% sequence similarity to the public database sequences (Fig. 2A). The always rare DNA also contained a group of common OTUs (96% similarity), but, notably, half the OTUs were uncommon, with only 84% identity to the public reference sequences (Fig. 2B). The abundant and always-rare RNA OTUs were common (98% and 96% sequence identity, respectively; Fig. 2) for the active fraction of the community, and the absence of uncommon OTUs in the rare 16S rRNA fraction indicates that the uncommon OTUs were never active. The low similarity to the SILVA database displayed by the uncommon rare OTUs (84% identity) suggests that they originated from undersampled ecosystems, not well covered by the public database. The closest relatives to the uncommon rare OTUs belonged to the Euryarchaeota

Deep Hydrothermal Vent Euryarchaeotic Group 6 (DHVEG-6), pMC1, and South African Gold Mine Euryarchaeotic Group-1 (SAGMEG-1) clusters, which are frequently detected in deep marine sediments (26). In contrast, the rare but common OTUs were identified as MGI and MGII.

Archaeal Community Structure, Dynamics, and Activity. The OTUs abundance followed a log-series distribution for the 16S rDNA and rRNA datasets (Fig. S2A and B), with most OTUs included in the first octaves (i.e., species characterized by a low number of reads). The abundant OTUs belonged mostly to MGI and MGII. A and MGII.B (Fig. S2C) but also to MGIII. In the active fraction (the 16S rRNA dataset), the major taxonomic groups (MGI, MGII.A, and MGII.B) represented $\sim 93\%$ of the reads (Fig. S2C). Interestingly, some abundant OTUs in the 16S rDNA dataset were less represented in the 16S rRNA dataset, for example, OTUs 2 and 9 affiliated with MGI, suggesting a weak activity. In contrast, some abundant OTUs were also very active, as shown by a greater relative abundance of 16S rRNA, such as OTU 28 affiliated with MGI (Fig. S2C).

The MGI sequences followed a seasonal pattern and were more abundant during winter (Fig. 3A). The MGI 16S rRNA dynamics showed the same trend as that for 16S rDNA, suggesting metabolically active communities. Our analysis showed that the MGI OTUs fell into four different clusters: A, B, C, and D (Fig. S3A). Most of the OTUs were affiliated with MGI.B, followed by MGI.A, which is closely related to *Nitrosopumilus maritimus*. These two clusters comprised all the abundant MGI OTUs. Interestingly, MGI.A and MGI.B sequences exhibited alternative patterns of 16S rDNA and rRNA representation. The MGI.A sequences outnumbered the MGI.B sequences in the 16S rDNA dataset (approximately two times more), whereas the opposite was observed for the 16S rRNA dataset (Fig. 4). This result suggested that MGI.B was much more active than MGI.A, even though it was not the most abundant in the ecosystem. The rare OTUs belonged to MGI.C, which is affiliated with sequences retrieved from deep waters (20) and distantly related to *Cenarchaeum symbiosum* and to MGI.D. This cluster was distinct from the others (89–92% similarity) and emerged earlier in the phylogeny. MGI.C was active when present, whereas MGI.D was not always active when present.

The MGII.A sequence abundance showed marked differences between seasons, with the highest relative abundance during the summer period (from May to October) and the lowest during the winter months (Fig. 3B). The 16S rRNA dataset revealed a similar seasonal pattern of activity. In contrast, MGII.B dominated in abundance and activity during winter, with the highest relative abundance in February and recurrent peaks each year (Fig. 3C). MGII.A was more active than MGII.B, consistent with its higher relative abundance (Fig. 3B and C). Euryarchaeota MGII.A was

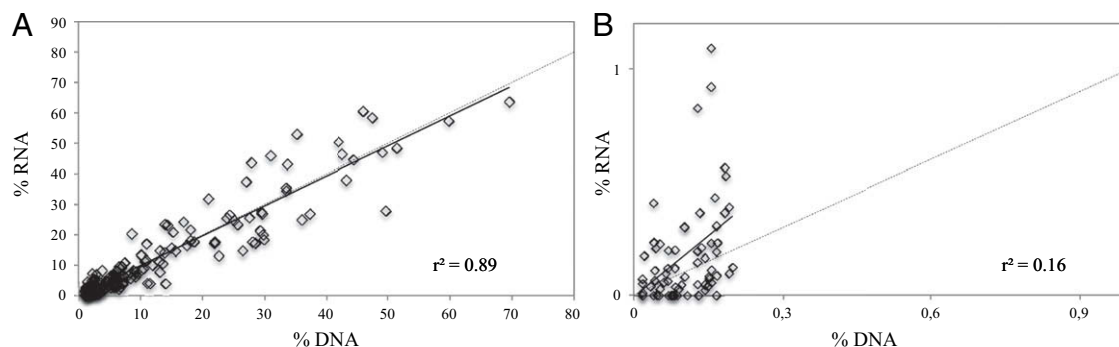


Fig. 1. 16S rDNA against 16S rRNA OTUs frequencies for abundant OTUs when they are abundant (A) and when they become rare (B). The RNA and DNA frequencies are plotted against each other for all abundant OTUs and all time points. The black line represents the regression, and the dotted line is the 1:1 line.

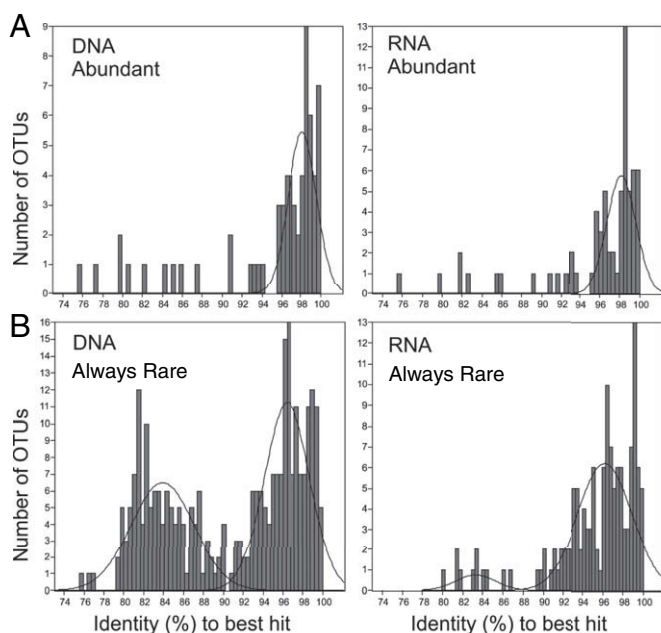


Fig. 2. Distribution of the percent identity from a comparison between public database sequences (SILVA) and abundant 16S rDNA and rRNA sequences (A) and always-rare 16S rDNA and rRNA sequences (B). The data are fitted to a model of normal distributions (black lines) that identifies groups of OTUs as common (i.e., a high percentage identity) or uncommon (i.e., a low percentage identity).

separated into two previously described (20) main subclusters (M and K; Fig. S4), and most of the sequences belonged to sub-cluster M, which was also the most active (Fig. S5A). The sub-cluster M activity pattern was different from that of sub-cluster K (Fig. S5). Most MGII.B sequences and activity were affiliated with the WHARN subcluster (Figs. S5B and S6) that corresponds to phylotypes II-CC, which are widely distributed in surface waters of various oceanic provinces (20). Other Euryarchaeota were affiliated with the MGIII and the RC-V cluster and with methanogenic lineages (Fig. S6).

Less abundant groups, including OTUs affiliated with MGIII, were also present and active during winter but were also detected in July 2008 and 2009, together with reduced activity. The Miscellaneous Euryarchaeotic Group (MEG) and DHVEG-6 did not present seasonal patterns of relative abundance and activity.

The canonical correspondence analysis plot (SI Materials and Methods) showed a clear difference between the activity of the two MGII clusters (Fig. S7): MGII.A appeared as a summer community associated mainly with temperature, whereas the activity of MGII.B was related to such winter features as nitrite, nitrate, and oxygen. These winter features also characterized the activity of MGI overall, and there were fewer differences between the different MGI clusters when considering the parameters followed in the present study. Contrary to MGII, MGI clusters were discriminated according to the second axis, which was positively correlated with phosphate (Fig. S7).

Discussion

Our long-term study of archaeal dynamics and activity in surface Mediterranean waters showed that rare Archaea were heterogeneous in their pattern of seasonal activity and phylogenetic affiliation. We propose that the rare archaeal biosphere could be divided into three different fractions classified as follows: the local seed bank, the nonlocal seed bank (or the alien colonizers), and the active-but-always-rare fraction.

The local seed bank represented Archaea that were rare but became abundant at certain times. When abundant, their 16S rDNA and 16S rRNA sequences were closely correlated, indicating that these OTUs were also active. Scatter plots of 16S rRNA vs. rDNA yielded an intercept at zero, suggesting that growth rates were constant as abundance varied (4). However, when these OTUs became rare, their 16S rDNA and rRNA sequences were poorly correlated, which, according to a described model (4), indicates increasing or decreasing growth rates as abundance decreases. Such variable activity suggests changing growth rates, possibly reflecting differences in the metabolic state of the cells as they cycle between abundant and rare fractions. Contrasting activity levels among rare microbes have been reported recently for Bacteria in a coastal system (4) and in lakes (7). Within the context of our seasonal study, the observations could indicate that these rare microorganisms are able to react to seasonal fluctuations of environmental conditions. Moreover, these Archaea could not be considered as being typically dormant cells because some of them lacked dormancy stages (i.e., were always active) and others had only short ones. We propose that this local seed bank maintains sufficient metabolic diversity to react to fluctuating environmental conditions.

The second fraction contained rare Archaea that were uncommon and always inactive in the northwestern Mediterranean Sea. They were aliens to the studied pelagic ecosystem, and their low similarity to database sequences indicates that they may originate from undersampled ecosystems, such as deep marine sediments. This nonlocal seed bank may be dispersed by such episodic events as river flooding, strong storms, or even atmospheric deposition. It is possible that these microorganisms may never grow in the water column as a result of a requirement for very different conditions to those found in the pelagic environment. This fraction of the rare archaeal biosphere could be on its way to extinction (13); alternatively, it may have the ability to

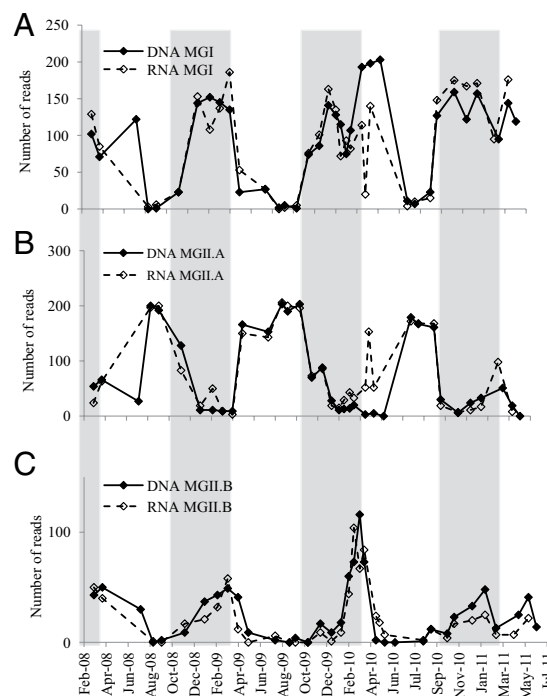


Fig. 3. Seasonal dynamics of the most abundant taxonomic groups in both the 16S rDNA and rRNA sequence datasets: (A) MGI, (B) MGII.A, and (C) MGII.B. Winter months are shown in gray; summer months are shown in white.

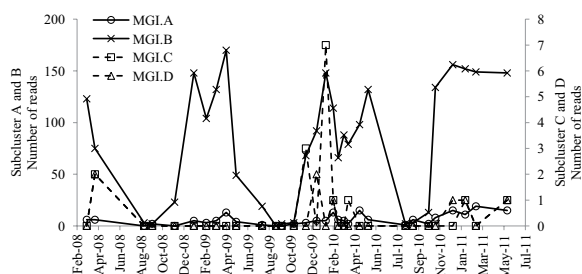


Fig. 4. Seasonal dynamics of active (16S rRNA dataset) MGI clusters: MGI.A, MGI.B, MGI.C, and MGI.D. Winter months are shown in gray; summer months are shown in white.

survive in a dormant stage representing a pool of potential colonizers of very different ecosystems.

Finally, the third fraction of rare Archaea was represented by cells that were always rare but actively growing, and their phylogenetic affiliation and seasonal dynamics were similar to those of the abundant Archaea. Although their activity suggests that these cells are not dormant, the fact that they never made the transition from rare to abundant indicates that they are not operating as a seed bank under the range of environmental conditions encountered during this study. Always rare but actively growing populations may be maintained at a low abundance because they are less efficient than others when competing for resources, or they may have a different life strategy that promotes rarity to minimize predation (11). Alternatively, these groups may be more susceptible to specific viral attack (27).

Although MGI was overall more abundant and active in winter, we found differences in activity between the clusters. Four MGI clusters were identified through our analysis: MGI.A (or I- α), MGI.B (or I- β), and MGI.C (I- γ) were first described in a large-scale study based on terminal restriction fragment length polymorphism (T-RFLP) (20), but only MGI.A and MGI.B were reported in a previous study conducted in the Mediterranean Sea (24). The similar delineation of MGI.A and MGI.B by cloning/sequencing (24) and pyrosequencing confirmed that shorter pyrosequencing reads may be as informative as near full-length sequences, and thus suitable for building phylogenies (28, 29). Interestingly, the MGI.A sequences were more abundant within the 16S rDNA fraction, whereas MGI.B was more abundant within the 16S rRNA fraction, suggesting that the most abundant MGI cluster (i.e., MGI.A) was not the most active. The rarest clusters, MGI.C and MGI.D, also presented different patterns of activity: MGI.C was active when present, yet MGI.D was not always active when present. Both of these rare clusters were active in winter, showing similar responses to such environmental conditions as oxygen, nitrite, nitrate, or Chlorophyll *a* (Chl*a*) content. However, there was a mismatch between their peaks of maximum activity. The variable activity levels for the different MGI clusters could suggest the presence of MGI clusters adapted to different niches. The clusters' recurrent seasonal activity patterns concomitant with reproducible environmental conditions may indicate an ecological specialization and predictable population-environment linkage, comforting the idea of separate ecotypes, as previously shown for Vibrionaceae populations (30). Although we could not precisely define niches in the present study, we hypothesize that the MGI clusters represent ecotypes that correspond to distinct metabolisms, as previously demonstrated for Bacteria (4). In fact, closely related SAR11 phylotypes were recently shown to have very different growth rates (4), which could correspond to differences in metabolism, such as those observed between the SAR11 ecotypes for phosphorus acquisition or glucose utilization (31, 32). Different metabolisms could thus allow

the MGI ecotypes to occupy a variety of niches and explain a global ecological success that is similar to that of SAR11 (33).

The MGI seasonal dynamics, illustrated by a higher relative abundance during winter, correlated with the seasonality of such surface water environmental parameters as nitrite and nitrate concentrations. MGI is thought to play an important role in the marine nitrogen cycle by oxidizing ammonia to nitrite (34, 35). In our survey, most of the 16S rDNA sequences belonged to MGI.A, which is closely related to the cultivated planktonic marine Archaea *N. maritimus* (36), an autotrophic ammonia-oxidizer that produces nitrite, suggesting the potential for ammonia oxidation of the MGI found in these waters. This hypothesis is supported by the MGI winter peaks that coincided with an increase in the nitrite and nitrate concentrations, followed by a decrease in ammonia. However, because phytoplankton can also release nitrite (37), the exact origin of winter nitrite in this ecosystem cannot be conclusively determined. Moreover, MGI.C was distantly related to *C. symbiosum*, which is able to use urea as energy and carbon source (38). As archaeal genes for urea utilization have been detected in the marine environment (39–41), we can also speculate that urea could be used as an alternative source of nitrogen and carbon by the MGI clusters that are not closely related to *N. maritimus*, which does not possess genes for urea utilization (42). The decrease in MGI abundance and activity was coincident with an increase in Chl_a, as previously shown in marine and freshwater ecosystems (43–45). This finding suggests two hypotheses: a limitation of MGI abundance by organic material excreted by phototrophic primary production (36) or a competition with phytoplankton for ammonium that may be unfavorable to MGI (46).

Euryarchaeota were present year-round, but the opposite seasonal dynamics of MGII.A and MGII.B suggests the presence of different ecotypes. The predominance of MGII.A in summer corresponds to a season of generally low abundance of MGII and Archaea in the northwestern Mediterranean Sea (24). The winter peak of MGII.B corresponds to the highest archaeal and MGI abundances (24), suggesting separate niches for the two MGII clusters in surface waters. Their abundance and activity dynamics might be affected by competition for resources with other organisms and could reflect the development of different strategies to improve their metabolic potential. For instance, genes encoding proteorhodopsin have been found in members of the Euryarchaeota MGII.A in the surface waters of the North Pacific (47), and recent studies showed a single copy of the proteorhodopsin (*pop*) gene in the reconstruction of a coastal MGII.A genome (48). In our study, the MGII.A sequences were only 90% similar to the proteorhodopsin clade from a previous work (47), but closer to the sequence reported in another (48) (96% similarity). We therefore hypothesize that phototrophic metabolism could be present in the Mediterranean Sea: MGII.A could use light as an energy source, explaining the summer peaks of abundance and activity, as irradiance is more important in this season. In contrast, the MGII.B distribution in the rRNA datasets correlated with nitrogen compounds, as with MGI.

The results based on the 16S rRNA/rDNA ratios could have been affected by the 16S rDNA copy number per genome. However, to our knowledge, all available complete genomes of mesophilic Archaea, including representatives from Euryarchaeota and Thaumarchaeota showed only one copy of 16S rDNA; thus, we assume that this is also the case in the natural communities. To assess how well our sequence data could represent the community abundance, we compared the pyrosequencing quantification to the metagenomic data obtained in September 2010 from the same sampling station through the Global Ocean Sampling project. The two methods showed a strong dominance of Euryarchaeota (95% and 88% for the pyrosequencing and metagenomic data, respectively) vs. Thaumarchaeota, suggesting no major primer bias at the phylum level in our approach. A

significant correlation has also been found between the abundance estimated by quantitative methods and pyrosequencing for Bacteria (4, 6, 9) and by sequencing approaches for MGI in the northwestern Mediterranean Sea (24). We therefore hypothesize that the relative sequence abundance measured in this study was comparable to the cell abundance dynamics.

In summary, this study clearly showed that the rare biosphere could not solely be characterized as a seed bank of dormant cells; rather, it is a complex association of indigenous and itinerant cell types with contrasted origins and fate that contribute to microbial interaction networks and metabolic processes in the environment. Our phylogenetic affiliation suggested that the diversity found within the environmental clusters of Archaea may correspond to different activity levels or growth rates, thus possibly illustrating different metabolism and life strategies. Our results show that we need to rethink our view of how abundant and rare microbes contribute to ecosystem processes.

Materials and Methods

Sampling and Environmental Parameters. Surface seawater (3 m) was collected monthly from March 2008 to June 2011 (40 samples) by using a 10-L Niskin bottle at the Service d'Observation du Laboratoire Arago station (42°31'N, 03°11'E) in the Bay of Banyuls-sur-Mer in France. The water was kept in 10-L high density polyethylene carboys in the dark until being processed in the laboratory (within 1.5 h). A subsample of 5 L was prefiltered through 3- μ m pore-size polycarbonate filters (Millipore), and the microbial biomass was collected on 0.22- μ m pore-size GV Sterivex cartridges (Millipore) and stored at -80 °C until nucleic acid extraction. The physicochemical parameters (Fig. S8) were provided by the Service d'Observation en Milieu Littoral (www.domino.u-bordeaux.fr/somlit_national).

The water sample used for the metagenomic analysis was collected at 3 m depth on 28 September 2010 as part of the J. Craig Venter Institute European Sampling Expedition following a protocol previously published (49). Annotation of the metagenomic data were performed through the J. Craig Venter Institute metagenomics analysis pipeline (San Diego) (50).

Nucleic Acid Extraction and Pyrosequencing. The nucleic acid extraction method was modified from Lami et al. (8) by using a combination of mechanical and enzymatic cell lysis applied directly to Sterivex cartridges, followed by extraction by using the AllPrep DNA/RNA kit (Qiagen). The RNA samples were tested for the presence of contaminating genomic DNA by PCR and then reverse-transcribed with random primers using the SuperScript III Reverse Transcriptase kit (Invitrogen). The amplification of the V3-V5 region of the 16S rRNA gene was performed by Research and Testing Laboratory (Lubbock, TX) with universal archaeal primers Arch349F (CCC TAC GGG GTG CAS CAG) and Arch806R (GGA CTA CVS GGG TAT CTA AT) (51), followed by pyrosequencing by using a Roche 454 GS-FLX system with titanium chemistry.

Bioinformatic Analysis and Statistics. The pyrosequencing data produced from the 80 samples (16S rDNA and 16S rRNA) represented 477,589 raw sequences. All sequences were checked against the following quality criteria: (i) no Ns; (ii) quality score ≥ 27 according to PANGAEA trimming (52); (iii) a minimum sequence length of 200 bp; (iv) no sequencing error in the forward primer; and (v) no chimeras [checked with UCHIME (53)]. The quality filtering step eliminated ~15% of all sequences (1.6% were chimeras). The remaining reads were clustered using USEARCH (54) at a 97% similarity threshold (55). For the taxonomic affiliation, we constructed a dedicated archaeal database

based on the SSURef 108 database of the SILVA project (56) and added annotated reference sequences from the Mediterranean Sea (24). The process was automated by PANAM (<http://code.google.com/p/panam-phylogenetic-annotation/downloads/list>) that constructs phylogenetic trees for taxonomic annotation (57) as detailed in *SI Materials and Methods*. After that step, all sequences affiliated to Bacteria were removed from the data set, leaving a total of 65,833 archaeal sequences for the 16S rDNA dataset and 52,181 sequences for the 16S rRNA dataset (Table S1). Phylogenetic trees containing only the main taxonomic groups detected by PANAM (MGI Thaumarchaeota, MGII.A and MGII.B Euryarchaeota), and environmental OTUs affiliated with those groups, are included as Figs. S3, S4, and S6.

For the analysis of the seasonal dynamics, the 16S rDNA and 16S rRNA samples were randomly resampled down to 208 sequences by using Daisy-Chopper (www.genomics.ceh.ac.uk/GeneSwyatch/). We chose to resample down to a relatively low number of sequences to retain the largest possible number of samples; a total of 12 samples were discarded because of a low number of sequences (< 208). However, for the analysis of the rare biosphere, a deeper sequencing effort was needed to define the rare Archaea, and only samples with >488 sequences were retained (55 samples). To verify if the different sampling cutoff could bias our analysis, we compared the seasonal dynamics based on 208 sequences per samples to that based on 488 sequences. The two results were similar for the major groups, as, for example, for MGI (Fig. S9). We also compared the number and identity of the abundant OTUs found for each cutoff. The entire 16S rDNA dataset, the one resampled at 208 sequences, and the one resampled at 488 sequences, showed 17, 18, and 21 abundant OTUs (> 1%), respectively (19, 22, and 21 for the 16S rRNA sequences). Notably, the abundant OTUs were always the same in the different datasets.

Defining Abundant and Rare Phylotypes. OTUs were considered abundant when they comprised more than 1% of the sequences (11) and were present in more than one sample. In contrast, rare OTUs were defined as OTUs representing $\leq 0.2\%$ of the sequences in a sample (present once in a sample of 488 sequences). This definition is well within the 0.1% to 1% range commonly considered (58), and is more strict than the 1% threshold used recently (4). OTUs were defined as always rare when they were rare in all the samples.

Representative sequences from all OTUs were compared with reference sequences from the entire SILVA database (56) using BlastN (59) to identify the percentage similarity between the queried sequences and their top hits. To assess the commonness of the sequences, the distribution of their percentage identity was plotted and fitted to normal distributions by using a maximum-likelihood method implemented in the mixture analysis of the PAST program (60). The method allowed us to define sequences as common (96–98% identity to database sequences) or uncommon (83% identity in average).

ACKNOWLEDGMENTS. We thank Cyrielle Tricoire, Eric Maria, and the captain and crew of the *Nereis II* for sample collection; and the people involved in the long-term series of hydrobiogeochemical data collected within the Service d'Observation en Milieu Littoral network (SOMLIT). This work was supported by a PhD fellowship from the French Ministère de l'Enseignement Supérieur et de la Recherche (to M.H.), a PhD fellowship from the French Conseil Régional d'Auvergne (to N.T.), and a CNRS Program Écosphère Continentale et Côtière (EC2CO, 2010–2012). The work of P.E.G. is supported by the Agence Nationale de la Recherche (ANR) project MICADO (ANR-11JSV7-003-01). J. Craig Venter Institute (JCVI) Global Ocean sampling, sequencing, and sequence analyses were funded by grants from the Beyster fund of the San Diego Foundation and the Life Technologies Foundation (to JCVI).

- Alonso-Saez L, et al. (2007) Seasonality in bacterial diversity in north-west Mediterranean coastal waters: Assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiol Ecol* 60:98–112.
- Mary I, et al. (2006) Seasonal dynamics of bacterioplankton community structure at a coastal station in the western English Channel. *Aquat Microb Ecol* 42: 119–126.
- Campbell BJ, Kirchman DL (2012) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J* 7:210–220.
- Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proc Natl Acad Sci USA* 108:12776–12781.
- Lennon JT, Jones SE (2011) Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nat Rev Microbiol* 9:119–130.
- Campbell BJ, Yu L, Straza TRA, Kirchman DL (2009) Temporal changes in bacterial rRNA and rRNA genes in Delaware (USA) coastal waters. *Aquat Microb Ecol* 57: 123–135.
- Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci USA* 107:5881–5886.
- Lami R, Ghiglione JF, Desvignes JF, West NJ, Lebaron P (2009) Annual patterns of presence and activity of marine bacteria monitored by 16S rDNA-16S rRNA fingerprints in the coastal NW Mediterranean Sea. *Aquat Microb Ecol* 54:199–210.
- Gaidos E, Rusch A, Ilardo M (2011) Ribosomal tag pyrosequencing of DNA and RNA from benthic coral reef microbiota: Community spatial structure, rare members and nitrogen-cycling guilds. *Environ Microbiol* 13:1138–1152.
- Sogin ML, et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* 103:12115–12120.
- Pedros-Alio C (2006) Marine microbial diversity: Can it be determined? *Trends Microbiol* 14:257–263.
- Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci USA* 106:22427–22432.
- Pedros-Alio C (2012) The rare bacterial biosphere. *Annu Rev Mar Sci* 4:449–466.

14. Francis CA, Beman JM, Kuypers MM (2007) New processes and players in the nitrogen cycle: The microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J* 1:19–27.
15. Hugoni M, et al. (2013) Dynamics of ammonia-oxidizing Archaea and Bacteria in contrasted freshwater ecosystems. *Res Microbiol*, 10.1016/j.resmic.2013.01.004.
16. Karner MB, DeLong EF, Karl DM (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409:507–510.
17. Herndl GJ, et al. (2005) Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* 71:2303–2309.
18. Teira E, et al. (2008) Linkages between bacterioplankton community composition, heterotrophic carbon cycling and environmental conditions in a highly dynamic coastal ecosystem. *Environ Microbiol* 10:906–917.
19. Varela MM, van Aken HM, Sintes E, Herndl GJ (2008) Latitudinal trends of Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern North Atlantic. *Environ Microbiol* 10:110–124.
20. Massana R, DeLong EF, Pedros-Alio C (2000) A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl Environ Microbiol* 66:1777–1787.
21. Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009) Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J* 3: 860–869.
22. Garcia-Martinez J, Rodriguez-Valera F (2000) Microdiversity of uncultured marine prokaryotes: The SAR11 cluster and the marine Archaea of Group I. *Environ Microbiol* 9:935–948.
23. Tully BJ, Nelson WC, Heidelberg JF (2012) Metagenomic analysis of a complex marine planktonic thaumarchaeal community from the Gulf of Maine. *Environ Microbiol* 14: 254–267.
24. Galand PE, Gutierrez-Provecho C, Massana R, Gasol J, Casamayor EO (2010) Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea. *Limnol Oceanogr* 55:2117–2125.
25. Koeppl A, et al. (2008) Identifying the fundamental units of bacterial diversity: A paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci USA* 105:2504–2509.
26. Teske A, Sorensen KB (2008) Uncultured archaea in deep marine subsurface sediments: Have we caught them all? *ISME J* 2:3–18.
27. Bouvier T, del Giorgio PA (2007) Key role of selective viral-induced mortality in determining marine bacterial community composition. *Environ Microbiol* 9:287–297.
28. Jeraldo P, Chia N, Goldenfeld N (2011) On the suitability of short reads of 16S rRNA for phylogeny-based analyses in environmental surveys. *Environ Microbiol* 13: 3000–3009.
29. Ragan-Kelley B, et al. (2012) Collaborative cloud-enabled tools allow rapid, reproducible biological insights. *ISME J* 7:461–464.
30. Szabo G, et al. (2013) Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J* 7:509–519.
31. Coleman ML, Chisholm SW (2010) Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* 107:18634–18639.
32. Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ (2010) The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* 12:490–500.
33. Brown MV, et al. (2012) Global biogeography of SAR11 marine bacteria. *Mol Syst Biol* 8:595.
34. Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* 102:14683–14688.
35. Wuchter C, et al. (2006) Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* 103:12317–12322.
36. Konneke M, et al. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437:543–546.
37. Lomas MW, Lipschultz F (2006) Forming the primary nitrite maximum: Nitrifiers or phytoplankton? *Limnol Oceanogr* 51:2453–2467.
38. Hallam SJ, et al. (2006) Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* 103:18296–18301.
39. Baker BJ, Lesniewski RA, Dick GJ (2012) Genome-enabled transcriptomics reveals archaeal populations that drive nitrification in a deep-sea hydrothermal plume. *ISME J* 6:2269–2279.
40. Konstantinidis KT, Bruff J, Karl DM, DeLong EF (2009) Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific subtropical gyre. *Appl Environ Microbiol* 75:5345–5355.
41. Yakimov MM, et al. (2011) Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary production in Tyrrhenian deep waters (Central Mediterranean Sea). *ISME J* 5:945–961.
42. Walker CB, et al. (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* 107:8818–8823.
43. Auguet JC, Casamayor EO (2008) A hotspot for cold crenarchaeota in the neuston of high mountain lakes. *Environ Microbiol* 10:1080–1086.
44. Herfort L, et al. (2007) Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiol Ecol* 62:242–257.
45. Robidart JC, et al. (2012) Seasonal *Synechococcus* and *Thaumarchaeal* population dynamics examined with high resolution with remote in situ instrumentation. *ISME J* 6:513–523.
46. Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* 461:976–979.
47. Frigaard NU, Martinez A, Mincer TJ, DeLong EF (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* 439:847–850.
48. Iverson V, et al. (2012) Untangling genomes from metagenomes: Revealing an uncultured class of marine *Euryarchaeota*. *Science* 335:587–590.
49. Rusch DB, et al. (2007) The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biol* 5:e77.
50. Tanenbaum DM, et al. (2010) The JCVI standard operating procedure for annotating prokaryotic metagenomic shotgun sequencing data. *Stand Genomic Sci* 2:229–237.
51. Takai K, Horikoshi K (2000) Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Appl Environ Microbiol* 66:5066–5072.
52. Giongo A, et al. (2010) PANGEA: Pipeline for analysis of next generation amplicons. *ISME J* 4:852–861.
53. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–2200.
54. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461.
55. Kim M, Morrison M, Yu Z (2011) Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* 84: 81–87.
56. Pruesse E, et al. (2007) SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35: 7188–7196.
57. Taib N, Mangot JF, Domaizon I, Bronner G, Debroas D (2013) Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: New insights into the freshwater protist diversity. *PLoS ONE*, 10.1371/journal.pone.0058950.
58. Fuhrman JA (2009) Microbial community structure and its functional implications. *Nature* 459:193–199.
59. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
60. Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. Available at <http://folk.uio.no/ohammer/past/>. Accessed May 20, 2009.

Supporting Information

Hugoni et al. 10.1073/pnas.1216863110

SI Materials and Methods

Nucleic Acid Extraction and Pyrosequencing. DNA and RNA were extracted by digesting cells directly in the Sterivex cartridge with lysis buffer (EDTA 40 mM; Tris 50 mM, pH 8.3, sucrose 0.75 M). Then heat/cold shocks treatments were performed, followed by the addition of lysozyme (40 mg·mL⁻¹) and incubation at 37 °C for at least 45 min. Then, proteinase K (20 mg·mL⁻¹) and SDS (0.2 g·mL⁻¹) were added, and the cartridges were incubated at 55 °C for 2 h. β -Mercaptoethanol was added to the cartridges content and nucleic acids extracted with the AllPrep DNA/RNA kit following the manufacturer's specifications (Qiagen). DNA and RNA yields were quantified by using a spectrophotometer (ND 1000; Nanodrop), and nucleic acid extracts were stored at -20 °C until analysis. RNA samples were tested for the presence of contaminant genomic DNA by PCR with archaeal 16S rDNA primers. Then, total RNA was reverse-transcribed with random primers by using the SuperScript III Reverse Transcriptase kit (Invitrogen) following the manufacturer's specifications. An identical set of reactions minus the reverse-transcriptase (i.e., no-RT reactions) were performed for each RNA extract; these reactions served as controls to examine the potential contributions of carryover genomic DNA on the PCR amplification of the cDNA. Nucleic acids were sent to Research and Testing Laboratory (Lubbock, TX) for amplification of the V3-V5 region of the 16S rRNA gene with universal archaeal primers Arch349F (CCC TAC GGG GTG CAS CAG) and Arch806R (GGA CTA CVS GGG TAT CTA AT) (1), followed by pyrosequencing on a Roche 454 GS-FLX system with titanium chemistry.

Pyrosequencing Data Analysis. Reference database. Sequences were compared with a dedicated database of reference sequences extracted from the SSURef 108 database from the SILVA project, which offers taxonomic information, quality assessment, and a curated alignment of SSU rRNA sequences (2). For the domain Archaea, the database includes 11,092 sequences with more than 900 bp, quality score >75%, and pintail value >50 according to the SILVA classification.

SILVA sequences together with annotated reference sequences from the Mediterranean Sea (3) were split into three monophyletic groups corresponding to the phyla Crenarchaeota, Thaumarchaeota, and Euryarchaeota, and a fourth group gathering sequences not affiliated to one of the three phyla. For each phyletic group, an outgroup containing one sequence from each of the other phyletic groups plus two distant sequences was added to the alignment to root the phyletic tree, and to specify the relatedness of early diverging sequences from the root of the group. Sequences from each phyletic group together with the outgroup sequences were retrieved from the SILVA alignment, and then trimmed to remove vertical gaps. A Hidden Markov Model (HMM) profile was built from each of the phyletic groups using HMMbuild from the HMMER package (4). A taxonomy file containing European Molecular Biology Laboratory taxonomy of each sequence from the reference database was also generated.

Sequence processing. First, a cleaning procedure was performed, whereby PANGEA functionalities (5) were used to remove short sequences (<200 bp), sequences with low quality score (≤ 27), sequences with at least one undetermined base, sequences with more than one mismatch with the forward primer, and to trim tags and adaptors. Then, we used UCHIME (6) for the detection of chimera. From the 477,589 raw sequences, 414,579 were kept after cleaning and 407,053 after chimera checking.

Second, clean reads were then clustered with UCLUST (7) at 97% identity.

Third, the operational taxonomic units (OTUs) were compared against the reference database with USEARCH (7). Then, following the taxonomy of its best hit, each sequence was appended to a phyletic group, together with its five best hits. The query sequences were sorted according to their assignment.

Fourth, homologous reads were aligned with the referenced sequences from the corresponding profile using HMMalign (4). Next, a phylogenetic tree was built for each phyletic profile using FASTTREE2 (8) with the Jukes-Cantor + Cat model and a bootstrap threshold of 100.

Fifth, trees were parsed to generate files containing the taxonomy of the inserted sequences. The taxonomy assessment was inferred by lowest common ancestor. All sequences affiliated with Bacteria or that were unclassified were discarded from further analysis.

Finally, the pipeline produces a file containing the monophyletic clusters with their bootstrap values, a list of all affiliated experimental sequences, their nearest reference neighbor and their taxonomy.

The package used for this analysis (named PANAM) can be obtained from <http://code.google.com/p/panam-phylogenetic-annotation/>. It comprises the reference sequences database, the taxonomy file, and reference profile alignments.

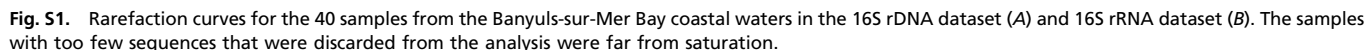
Statistical Analyses. Canonical correspondence analysis (CCA) was performed in XL Stat 2010 (Addinsoft) to assess the relationships between taxonomic groups and environmental parameters. Reads for each taxonomic cluster considered were pooled for each sampling date. CCA was performed on 10 environmental factors (temperature, salinity, oxygen, Chl *a* content, pH, nitrite, nitrate, phosphate, silicate, and ammonium concentrations) with the taxonomic cluster abundance (inferred from read number) matrix of 16S rRNA dataset.

Kendall correlations were calculated to evaluate potential significant relationships between 16S rRNA and rDNA frequency for each OTU and time point (Fig. 1). Correlations were considered significant when $P < 0.05$.

SI Results

Environmental Parameters. Analyses of 3.5 y of environmental data from the coastal surface waters (3 m) of Banyuls-sur-Mer (France) indicated concentrations of inorganic nutrients and phytoplankton biomass characteristic of an oligotrophic environment during most of the year (Fig. S8). Temperature peaked in the late summer; on the contrary, oxygen concentrations were highest during winter. Some parameters displayed a consistent temporal pattern with highest abundance in winter, i.e., nitrate, silicate, or nitrite, and, to a lesser extent, Chl *a* (Fig. S8), illustrating the hydrodynamic features of the surface waters studied, particularly during winter and spring when precipitation and freshwater input from the land, and sediment resuspension after heavy storms, are recurrent events.

Comparison with Metagenomics Data. To assess possible bias, we compared pyrosequencing vs. metagenomics sequence counts from the J. Craig Venter Institute metagenomic analysis. The proportion of Euryarchaeota vs. Thaumarchaeota obtained by pyrosequencing for the September 28, 2010, sample was similar to the one obtained through the J. Craig Venter Institute metagenomic analysis (95% and 88% Euryarchaeota for pyrosequencing and metagenomics, respectively).



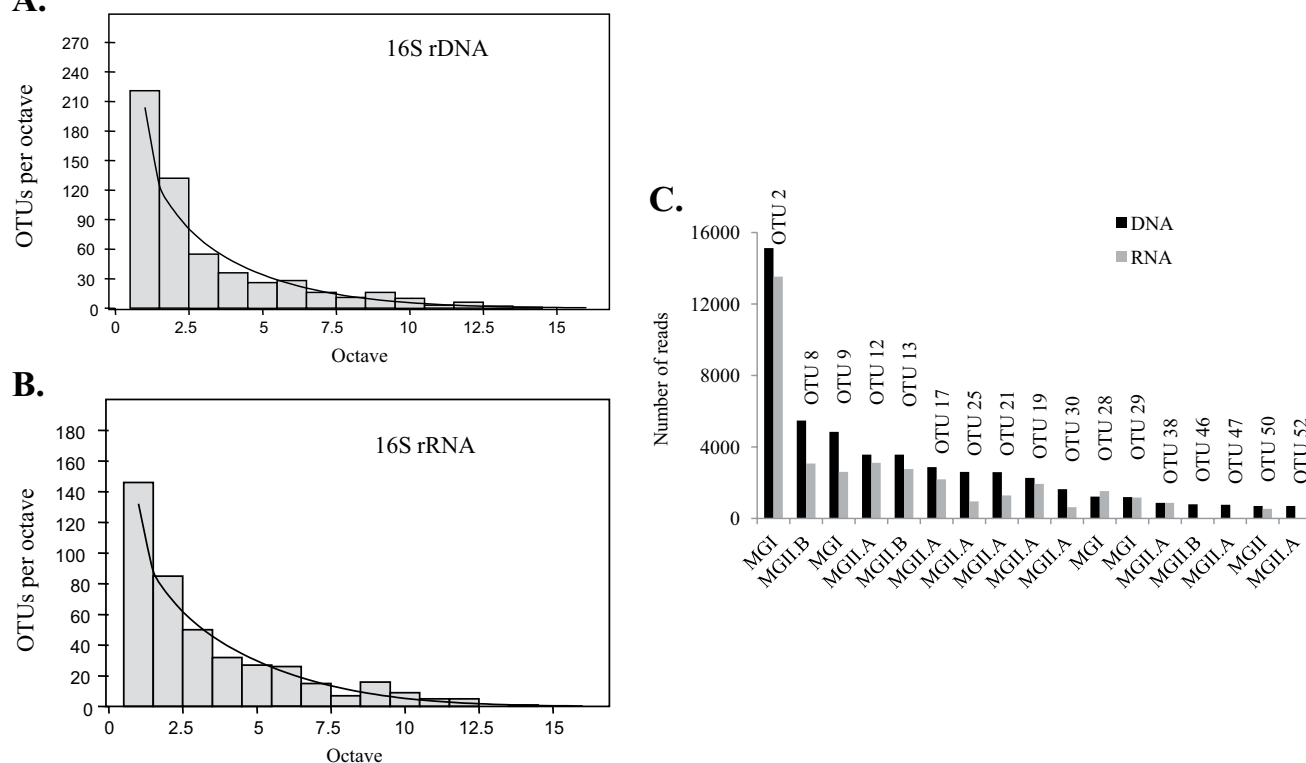


Fig. S2. Abundance distribution of archaeal OTUs obtained from the 40 samples from March 2008 to June 2011 at the Service d'Observation du Laboratoire Arago station in the 16S rDNA dataset (A) and the 16S rRNA dataset (B). The abundance models predicting the frequency of each abundance class are shown as lines. In both datasets, OTUs abundances were predicted by a log-series model. Octaves refer to power-of-two abundance classes. (C) The curve represents a detailed view of the respective abundant OTUs found in both 16S rDNA and 16S rRNA datasets.

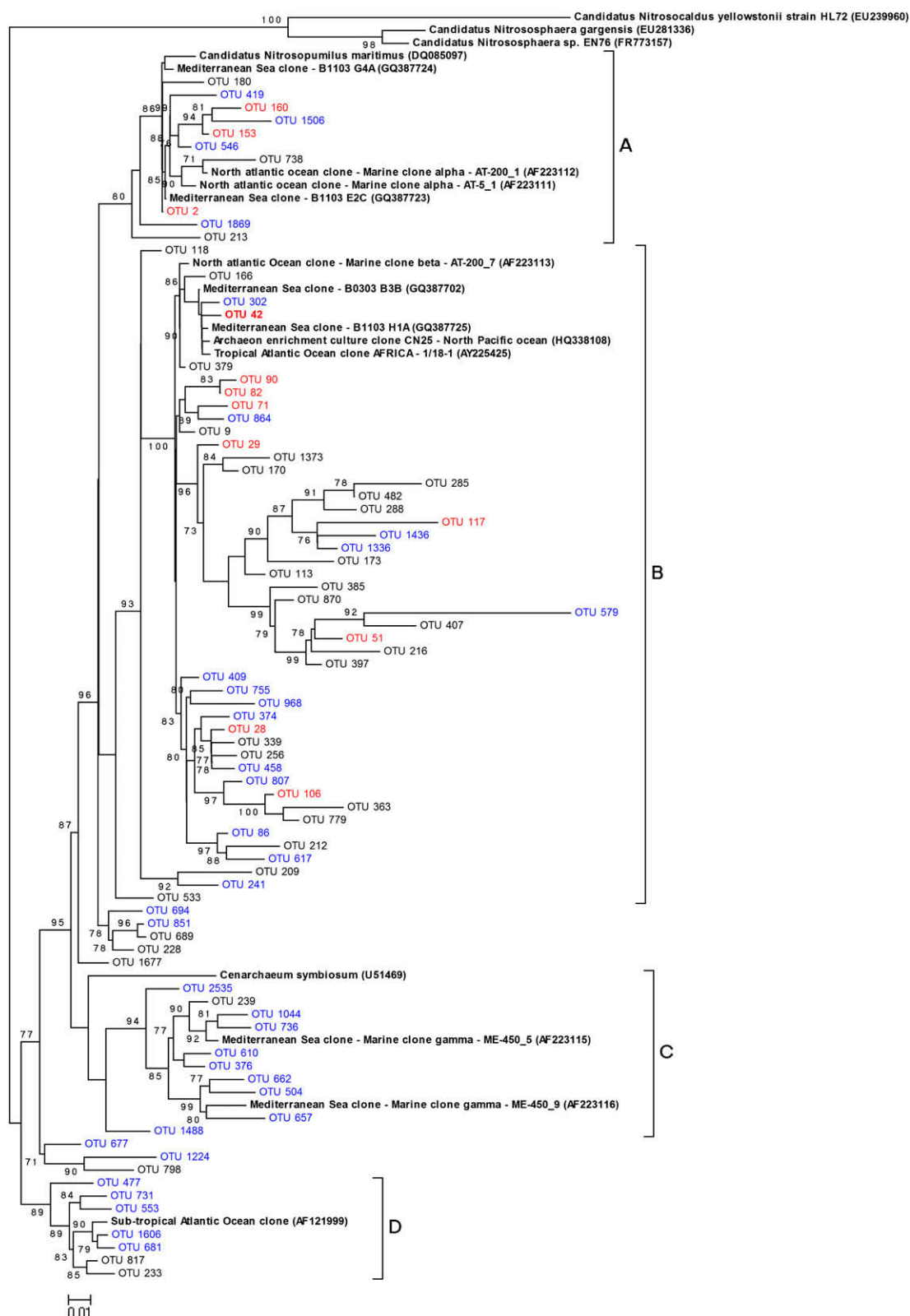


Fig. S3. Phylogenetic tree representing the position of marine group (MG) I 16S rDNA sequences from the Banyuls-sur-Mer Microbial Observatory. Reference sequences retrieved from GenBank are in bold. Abundant OTUs are represented in red, and always-rare ones are in blue. Bootstrap values >70 are shown expressed as a percentage of 100 replicates. (Scale bar: 10% sequence divergence.)

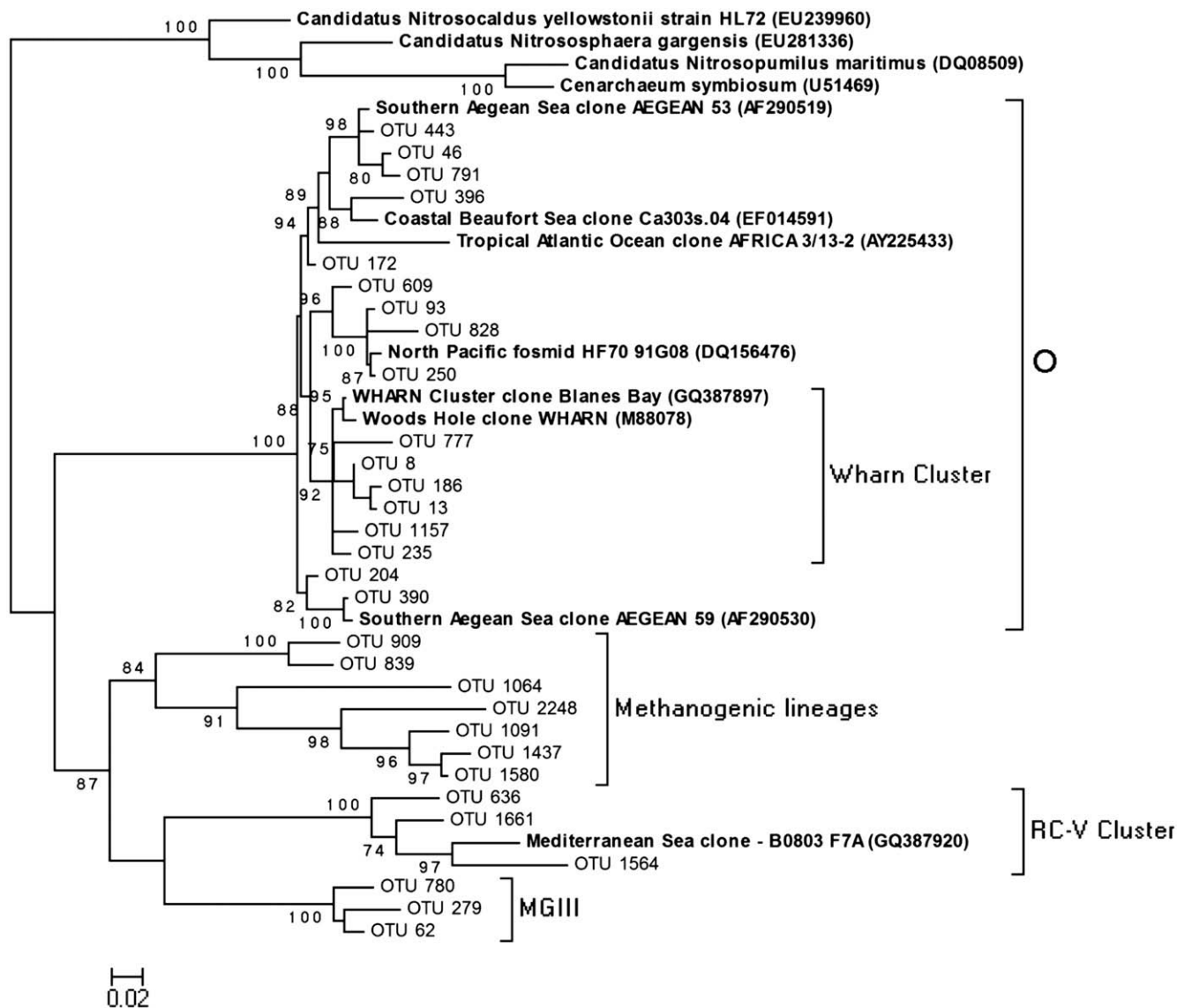
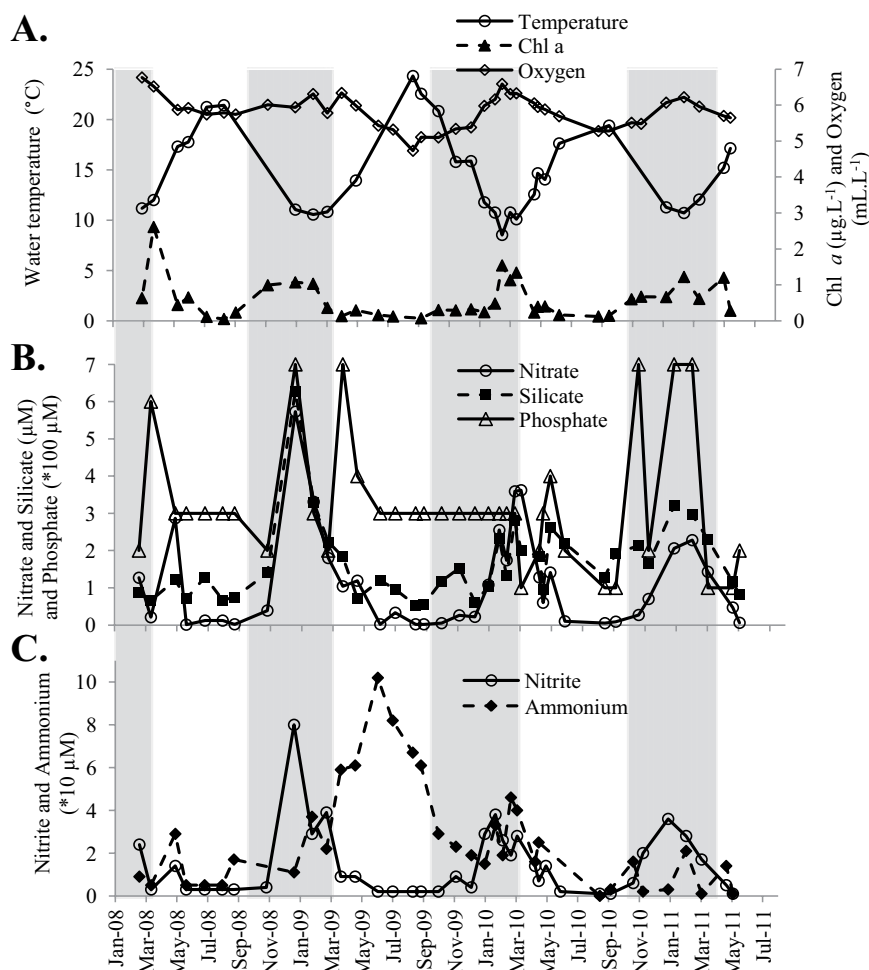


Fig. S6. Phylogenetic tree representing the position of MGII.B 16S rDNA sequences from the Banyuls-sur-Mer Microbial Observatory. Reference sequences retrieved from GenBank are in bold. Abundant OTUs are represented in red, and always-rare ones are in blue. Bootstrap values >70 are shown expressed as a percentage of 100 replicates. (Scale bar: 20% sequence divergence.)



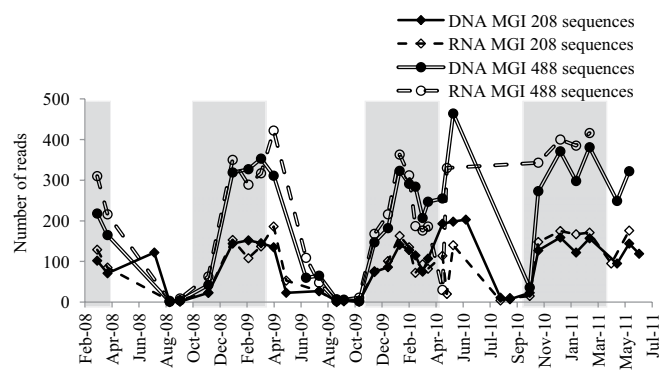


Fig. S9. Seasonal dynamics of MGI in both 16S rDNA and rRNA sequence datasets for the two normalization threshold 288 and 488 sequences. Winter months are shown in grey; summer months are shown in white.

PNAS

Environmental parameters (temperature, salinity, oxygen, pH, ammonium, nitrate, nitrite, phosphate, silicate, and Chla) associated to each point are presented. ND, not determined; QC, quality-checked; SOLA, Service d'Observation du Laboratoire Arago.

4.3 Article 2 : New insights into active archaeal groups in oxygen depleted zones of lacustrine ecosystems

Mylène Hugoni, Isabelle Domaizon, Najwa Taib, Gisèle Bronner, Corinne Biderre-Petit,
Hélène Agogué, Pierre E. Galand, Didier Debroas, Isabelle Mary
En préparation

New insights into active archaeal groups in oxygen depleted zones of lacustrine ecosystems

Mylène Hugoni^{1,2}, Isabelle Domaizon³, Najwa Taib^{1,2}, Gisèle Bronner^{1,2}, Corinne Biderre-
5 Petit^{1,2}, Hélène Agogué⁴, Pierre E. Galand^{5,6}, Didier Debroas^{1,2}, Isabelle Mary^{1,2}

(1) Clermont Université, Université Blaise Pascal, Laboratoire "Microorganismes : Génome et
Environnement", BP 10448, F-63000 Clermont-Ferrand, France;

(2) CNRS, UMR 6023, LMGE, F-63171 Aubière, France;

10 (3) Institut National de la Recherche Agronomique, UMR 42 Centre Alpin de Recherche sur
les Réseaux Trophiques et Ecosystèmes Limniques, F-74200 Thonon les Bains, France;

(4) Littoral, Environnement et Sociétés, UMR 7266, CNRS, University of La Rochelle, 17000
La Rochelle, France;

(5) Université Pierre et Marie Curie–Paris 6, UMR 8222, Laboratoire d’Ecogéochimie des
15 Environments Benthiques (LECOB), UMR 7621, Laboratoire d’Océanographie Microbienne
(LOMIC), Observatoire Océanologique, F-66650 Banyuls-sur-Mer, France;

(6) CNRS, UMR 8222, LECOBI, Observatoire Océanologique, F-66650 Banyuls-sur-Mer,
France.

20 **Correspondence:** Didier Debroas, LMGE, Laboratoire Microorganismes: Génome et
Environnement, UMR CNRS 6023, Université Blaise Pascal (Clermont-Ferrand II), 24
avenue des Landais, BP 80026, Aubière 63171, France. Tel.: + 33 4 73 40 78 37; fax: + 33 4
73 40 76 70; e-mail: didier.debroas@univ-bpclermont.fr

Running title: Active archaeal communities in lacustrine ecosystems

Abstract

Meromictic lakes are of specific interests in studying microbial assemblages down stratified water column. While aquatic *Archaea* are important players among microbial plankton and significantly contribute to biogeochemical cycles, details regarding their activity are still scarce. Active archaeal community structure (i.e. 16S rRNA) was investigated monthly during a year through two contrasted zones, the oxycline and the anoxic zone of two deep lakes, one presenting a permanent anoxic zone (Lake Pavin) while mixing of the entire water column in the second could occur (Lake Bourget). Different patterns in active archaeal distribution in the two lacustrine systems were evidenced, with one clearly dominated by *Thaumarchaeota* MGI, while the other was characterized by great temporal rearrangements of euryarchaeal phylotypes that dominated. The microbial rearrangements in the monimolimnion of meromictic lakes were therefore more dynamic than previously described. In those lakes, thaumarchaeal activity in ammonia oxidation (i.e. *amoA* transcripts) was lower than expected (based on the 16S rRNA dataset), suggesting alternative metabolic pathways for these microorganisms. In addition, some groups such as the crenarchaeal group C3 or even the Miscellaneous Euryarchaeotic Group (MEG) presented a high activity in those ecosystems, suggesting a role to define in ecosystem functioning.

Introduction

Planktonic freshwater habitats have emerged as one of the largest reservoirs of archaeal diversity (Auguet *et al.*, 2011; Lliros *et al.*, 2010). Indeed, reported natural abundances of uncultured *Archaea* range between 1 and 20% of total bacterioplankton in freshwater ecosystems (Glockner *et al.*, 1999; Keough *et al.*, 2003; Pernthaler *et al.*, 1998). In lacustrine ecosystems, *Archaea* are involved in two main biogeochemical cycles: the nitrogen and carbon metabolism. Indeed, recovery of archaeal genes encoding ammonia monooxygenase

subunit A (*amoA*), within *Thaumarchaeota* Marine Group I, in oxic layers of freshwater ecosystems (Auguet *et al.*, 2011; Pouliot *et al.*, 2009) had led to the belief that they could oxidize ammonia to nitrite, which was earlier thought to be achieved only by Ammonia Oxidizing *Bacteria* (AOB, (Kowalchuk and Stephen, 2001)). This was further confirmed by
55 microcosm experiments (Treusch *et al.*, 2005) and cultivation of several Ammonia Oxidizing *Archaea* (AOA) like *Nitrosopumilus maritimus* (Konneke *et al.*, 2005) or *Nitrososphaera gargensis* species (Hatzenpichler *et al.*, 2008). On the other hand, in lacustrine deeper layers (i.e. anoxic zone and sediments), methanogenesis performed by *Euryarchaeota*, is the most important terminal processes in the anaerobic degradation of organic material (Glissman *et*
60 *al.*, 2004). Initially thought to happen in anoxic environments, this process was recently described in a well-oxygenated water column of an oligotrophic lake (Grossart *et al.*, 2011). Therefore, *Euryarchaeota* could also be important players of the carbon cycling in oxic aquatic environments and water to air methane flux (Grossart *et al.*, 2011).

Based on their physiology, oxygen is a key factor structuring the archaeal
65 communities. In this way, meromictic lakes appeared as excellent models for studying microbial changes along a redox potential from depleted oxygen layers to anoxic ones. Moreover, the steep vertical physicochemical gradients retrieved in those ecosystems could provide optimal niches for microbial growth and differentiation (Pouliot *et al.*, 2009). While several studies focused on meromictic lakes stratified by salinity (Bosshard *et al.*, 2000;
70 Comeau *et al.*, 2012; Pouliot *et al.*, 2009), few investigated lakes that were meromictic because of their depth and/or their shape (Lehours *et al.*, 2005).

The distribution of archaeal metabolisms corresponded obviously to a vertical structuration of archaeal communities down the water column depending on physico-chemical conditions. *Thaumarchaeota* tend to dominate in oxycline and halocline as described in Lake
75 Kivu (Lliros *et al.*, 2010) and in an Arctic saline lake (Comeau *et al.*, 2012); while in deeper

waters (i.e. anoxic hypolimnion) various groups are retrieved as the highly diverse Miscellaneous Crenarchaeotic Group (MCG, (Inagaki *et al.*, 2003)), methanogenic lineages and uncultured Marine Benthic Group D (MBG-D, (Galand *et al.*, 2012)), or the uncultured crenarchaeal group C3 (Comeau *et al.*, 2012). In those lacustrine water layers, *Euryarchaeota* 80 affiliated with the Lake Dagow Sediment cluster (LDS (Glissman *et al.*, 2004)) and the RC-V cluster (Rice Cluster-V (Großkopf *et al.*, 1998)) were also retrieved. Those groups were highly diverse and frequently retrieved in lakes (Glissman *et al.*, 2004; Jurgens, 2000) and rivers (Galand *et al.*, 2006; Herfort *et al.*, 2009), suggesting that they have a key functional role in freshwater habitats. Those results showed a clear vertical segregation above and below 85 the oxycline zone, and further investigations on active populations are necessary.

While most of these studies investigating lacustrine ecosystems focused on archaeal diversity and quantification of gene abundance, few stressed on active archaeal assemblages (La Cono *et al.*, 2013; Vissers *et al.*, 2013). However, recent studies have shown the importance of differentiating the active communities from the total communities (Campbell *et* 90 *al.*, 2011; Jones and Lennon, 2010) and one method to explore an aspect of activity (i.e. the growth rate for specific taxa) is to investigate microbial communities with both 16S rDNA and 16S rRNA (Campbell *et al.*, 2009; Lami *et al.*, 2009).

Moreover, as emphasized by Hatzenpichler and coworkers (2012), active AOA have rarely been studied in freshwater ecosystems (Hatzenpichler, 2012). In the same way, 95 diversity analysis of archaeal assemblages through the water column were mostly conducted at specific times (Lehours *et al.*, 2005) or during contrasted period, as for example the summer and winter periods (Auguet *et al.*, 2011; Lliros *et al.*, 2008).

Here, two deep freshwater ecosystems were investigated to test whether different niches exist for active archaeal communities. The community structure of active *Archaea* was 100 characterized over one year in Lake Pavin and Bourget, by pyrosequencing 16S rDNA and

rRNA in the V4-V5 region. We stressed on two different zones controlling archaeal activity: the oxycline and the anoxic zone. Indeed, archaeal abundance and metabolic abilities to oxidize ammonia in the oxycline and the methanogenesis process mediated by *Euryarchaeota* in the anoxic zone, stressed the importance of such contrasted zones on archaeal activities, and could thus represent specific niches for archaeal groups with unknown metabolic pathways. This was supported by the weak abundance of AOA *amoA* transcripts in both lakes, evaluated through quantitative PCR during the whole year.

Materials and methods

Study sites, sampling and environmental parameters

This study was conducted on two lakes located in France: Lake Pavin in the Massif Central, and Lake Bourget at the edge of the Alps. Lake Pavin (45°55'N; 2°54'E) is a meromictic, oligomesotrophic freshwater lake, situated at an altitude of 1197 m, with a maximal depth of 92 m. It is fed by atmospheric precipitations and numerous superficial and sub-lacustrine springs (Viollier *et al.*, 1997). Lake Bourget (45°44'N; 5°51'E), situated at an altitude of 231 m, is classified as a meromictic lake but mixing events in the whole water column (maximal depth of 145.5m) has occurred frequently in the recent years (observatory on deep peri-alpine lakes, INRA Thonon Les Bains). It presented an oxic-anoxic transition zone and a deeper layer, which is anoxic during several months per year.

Water samples were collected monthly, during 2011, in the oxycline (located at 45 m and 130 m for Lake Pavin and Bourget respectively, during thermal stratification), and in the anoxic zone (80 m and 140 m for Lake Pavin and Bourget respectively) using a Van Dorn bottle at a permanent station located at the deepest zone of the water column. Water temperature and dissolved oxygen content were determined with a multiparameter probe (YSI GRANT 3800). Phosphorus (P-PO₄), nitrate (N-NO₃) and ammonium (N-NH₄) contents were

analyzed using standard American Public Health Association 100 (1992) methods. Chlorophyll-*a* (Chl*a*) content was determined by spectrophotometry (Lorenzen, 1967a; Strickland and Parsons, 1968).

130 **Nucleic acids extraction and pyrosequencing**

A sub-sample of water (300 mL), added with an equal volume of RNA Later (ammonia sulfate 7.93M, sodium citrate 0.025M, EDTA 0.02M qsp 1.5L of RNase free water, pH 5.2), was pre-filtered through 5- μ m pore-size polycarbonate filters (Millipore) and collected on 0.2- μ m pore-size (pressure <10 kPa) polycarbonate filters (Millipore) before storage at -80°C
135 until nucleic acid extraction. The nucleic acids extraction method was modified from (Hugoni *et al.*, 2013b) using a combination of mechanic and enzymatic cell lysis, followed by the extraction using the AllPrep DNA/RNA kit (Qiagen, Valencia, CA). The RNA samples were tested for the presence of contaminating genomic DNA by PCR and then reverse transcribed with random primers using the SuperScript[®] VILO (Invitrogen). The amplification of the V4-
140 V5 region of the 16S rDNA and rRNA was performed with universal archaeal primers Arch519F (Herfort *et al.*, 2009) and Arch915R (Casamayor *et al.*, 2002) (Supplementary Table 1). Pyrosequencing was realized by the GINA Plateform (Clermont-Ferrand, France), using a Roche 454 GS-FLX system with titanium chemistry.

145 **Bioinformatic and statistical analyses**

The pyrosequencing data for both 16S rDNA and 16S rRNA datasets represented 698,901 raw sequences. All the sequences were checked against the following quality criteria: (i) no Ns; (ii) quality score ≥ 25 ; (iii) a minimum sequence length of 200 bp; (iv) and no sequencing error in the forward primer. The remaining reads were clustered at a 97% similarity threshold
150 and representative OTUs were inserted in phylogenetic trees for taxonomic annotation. This

process was automated by PANAM (<http://code.google.com/p/panam-phylogenetic-annotation/downloads/list>) (Taib *et al.*, 2013). After the removal of sequences affiliated with *Bacteria*, the dataset contained a total of 104,675 archaeal sequences for the 16S rDNA dataset and 117,913 sequences for the 16S rRNA dataset (Supplementary Table 2). To
 155 compare 16SrDNA and 16SrRNA, samples were then randomly resampled down to 223 and 1247 sequences for Lake Pavin and Bourget, respectively.

A Venn diagram was generated with Venny (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) to evaluate common OTUs between the lakes and the layers considered.

To explain the variation of archaeal communities structure, redundancy analysis
 160 (RDA) was used after a forward selection (Borcard *et al.*, 1992) of the environmental variables (temperature, oxygen, phosphate, and ammonia concentrations) susceptible to explain a significant part of changes in archaeal taxonomic clusters abundance (inferred from reads number). This analysis was performed with R associated to the package VEGAN (<http://cran.r-project.org/web/packages/vegan/index.html>).

Quantitative PCR analysis

The qPCR protocol was modified from (Hugoni *et al.*, 2013a) and used to quantify archaeal 16S rDNA genes on both DNA and cDNA, and bacterial and thaumarchaeal *amoA* transcripts on cDNA. The reaction mixture (25 µL) contained MESA GREEN qPCR MasterMix Plus for
 170 SYBR Assay[®] (1X, Eurogentec) added with 0.8 µg of BSA, 0.7 µM of primers (Supplementary Table 1) and ultra-pure sterile water. One µL of nucleic acids was added to 24µL of mix in each well. All qPCR reactions were triplicated and consisted of an initial denaturing step at 94°C (for 15min for thaumarchaeal *amoA*, and 5min for archaeal 16S rDNA and bacterial *amoA* genes) and followed by (thaumarchaeal *amoA*: 94°C 15sec, 52°C
 175 30sec, 72°C 30sec; bacterial *amoA*: 95°C 30sec, 56°C 40sec, 72°C 2min; archaeal 16S rDNA:

94°C 30sec, 57°C 40sec, 72°C 40sec). Standard curves were generated from mix of clone's representative of the environments studied. All reactions were performed with standard curves spanning from 10^1 to 10^8 copies per μL . Mean PCR efficiencies and correlation coefficients for standard curves were as follows: for the thaumarchaeal *amoA* assay, 108 %, $r^2 = 1.00$, for the archaeal 16S rDNA assay, 104 %, $r^2 = 0.8$, and for the bacterial *amoA* assay, 107 %, $r^2 = 1.00$.

Results

Physico-chemical and biological characteristics of the lakes

Lake Pavin had a permanent oxycline and an anoxic zone while Lake Bourget only had temporary oxycline and anoxic conditions in the bottom of the lake (Supplementary Table 3). During the whole sampling year, the oxygen concentration ranging from 1 and 2.89 mg.L^{-1} in the oxycline of Lake Pavin (45 m). In contrast, Lake Bourget was characterized by an oxycline zone at 130 m, oxic from March to June (oxygen ranging from 7.93 and 9.47 mg.L^{-1}) and depleted of oxygen from July to December (oxygen ranging from 3.24 and 6.61 mg.L^{-1}). In the same way, at 140 m, the lake was oxygenated from March to June (oxygen ranging from 5.96 and 9.37 mg.L^{-1}), and almost anoxic from August to December (oxygen ranging from 0.06 and 1.34 mg.L^{-1}). Interestingly, in Lake Bourget, average ammonia concentrations were very low compared to Lake Pavin while the contrary was observed for nitrates content (Supplementary Table 3).

Abundance of active *Archaea*

In Lake Pavin, archaeal abundance (i.e. 16S rDNA (Supplementary Figure 1.A)) and activity (i.e. 16S rRNA (Figure 1.A)) were less important in the oxycline than in the anoxic zone. In Lake Bourget there were less difference between the two water layers for both archaeal

abundance and activity (Figure 1.B and Supplementary Figure 1.B). For both lakes and water layers, clear temporal changes were observed: in the oxycline of Lake Pavin, *Archaea* were more active during spring and autumn while in the anoxic zone, their maximal activity was seen during late spring and summer. In Lake Bourget an increase of active *Archaea* abundance was observed during the winter period in both the oxycline zone and the anoxic one, reaching respectively 14,000 and 60,000 transcripts mL⁻¹ in January (Figure 1.B).

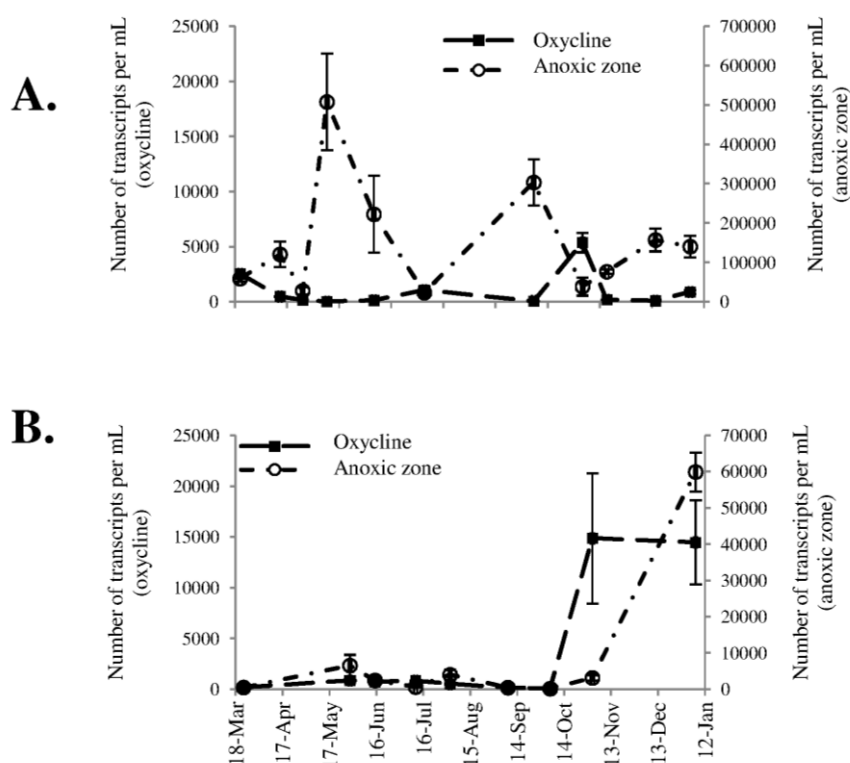


Figure 1. Abundances of 16S rRNA per mL in Lake Pavin (A) and Lake Bourget (B). Transcripts copy numbers were quantified by qPCR over the course of a whole year (2011) among two sampling depths, 45 m (oxycline) and 80 m (anoxic zone) for Lake Pavin and 130 m (oxycline) and 140 m (anoxic zone) for Lake Bourget.

Vertical segregation of archaeal communities

Pyrosequencing was used to evaluate archaeal community structure in contrasted zones (oxycline and anoxic zone) of freshwater ecosystems. The OTUs distributions were compared between lakes and between the different layers by pooling all sampling dates for both 16S rDNA and 16S rRNA datasets (Figure 2). Interestingly, only 5 OTUs were exclusively

retrieved in the anoxic zone of both lakes investigated. These OTUs were affiliated to *Euryarchaeota* and more precisely to MEG (2 OTUs), LDS (2 OTUs) and DHVEG-6 (1 OTU). Surprisingly, in Lake Bourget, 27.7% of the total OTUs were common in the oxycline and the anoxic zone; while in Lake Pavin 32.7% of the OTUs were common in the two different zones. In contrast, 20 OTUs were exclusively retrieved in the oxycline zone of both lakes (Figure 2). Among those OTUs, eight were affiliated with MGI, nine with LDS, and one in RC-V, DHVEG-6 and *Thermoplasmatales* AMOS1A-4113-D04 groups.

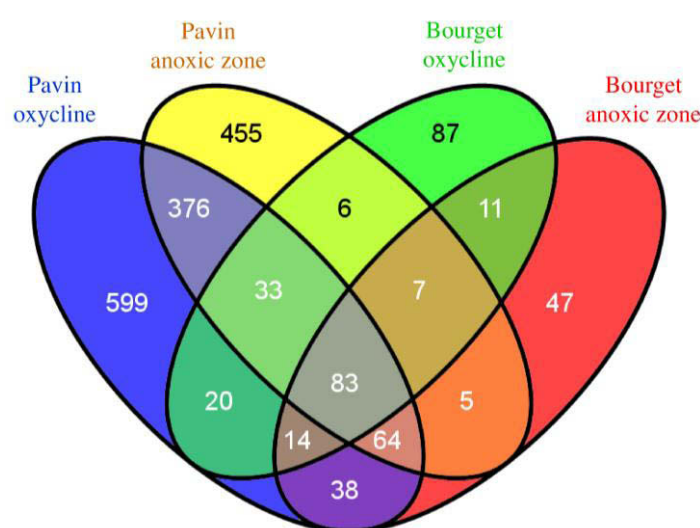


Figure 2. Venn diagram showing the number of unique and shared OTUs for the oxycline and the anoxic zone of both lakes.

Samples were then randomly resampled to compare 16SrDNA and 16SrRNA. Average reads counts for 16S rDNA and 16S rRNA datasets were calculated for major taxonomic groups retrieved in each ecosystem. More archaeal clades were retrieved from Lake Pavin compared to Lake Bourget for both the 16S rDNA and 16S rRNA fractions (Figure 3). In the oxycline of this lake, some groups were abundant and not very active (16SrRNA/16SrDNA ratios <1) like MGI or LDS, while others, such as MEG or *Methanosaeta*, had a greater activity in this zone (Figure 3). The same active groups (ratio >1) were retrieved in the anoxic zone, except for MCG that was replaced by active uncultured *Methanomicrobiales* in this zone. Lake Bourget was characterized by the dominance of MGI

that was active in this ecosystem (Figure 3). Nevertheless, the activity of this group is much more important in the oxycline while other groups like LDS or MCG were active in the anoxic zone. These results clearly highlighted a contrasted picture between dominant and active *Archaea*.

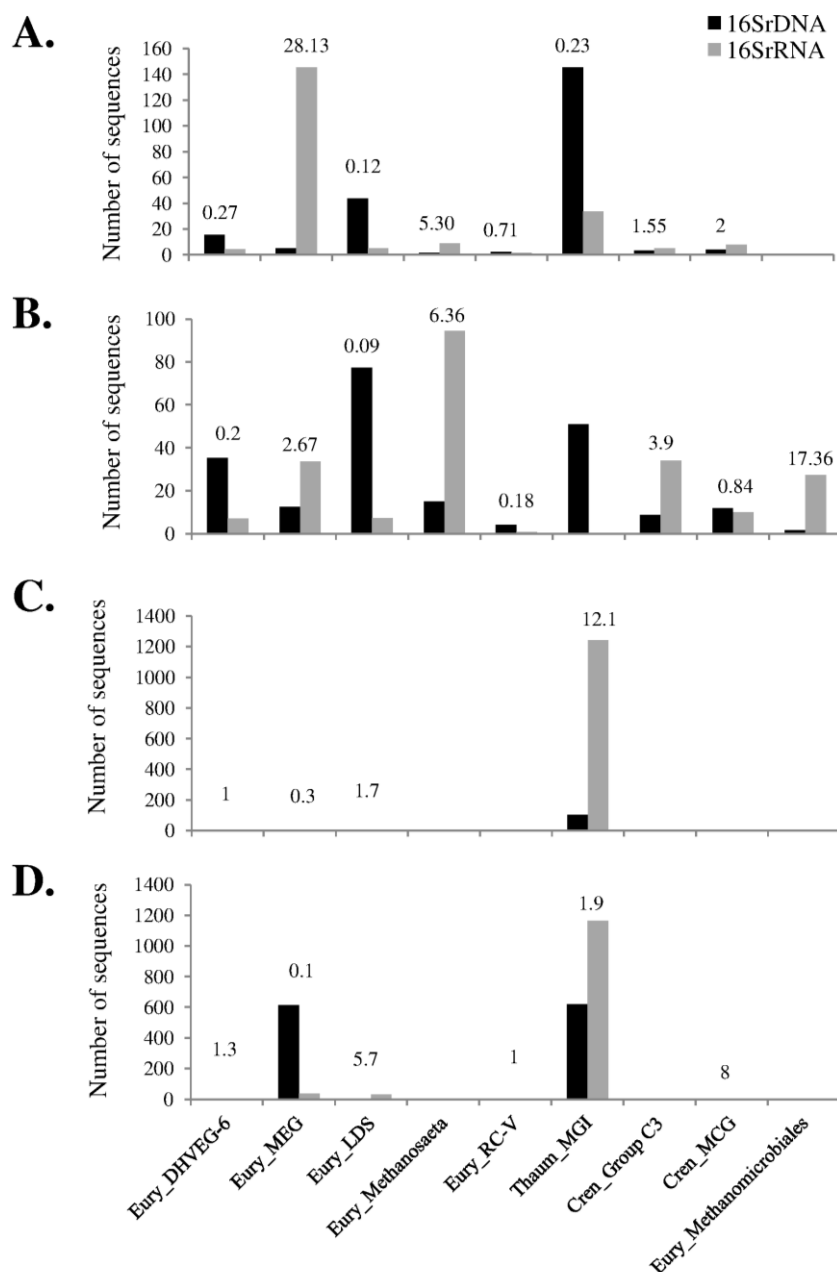


Figure 3. Number of sequences affiliated with major taxonomic groups retrieved in both 16S rDNA and 16S rRNA datasets of the oxycline (A) and the anoxic zone (B) of Lake Pavin, and in the oxycline (C) and the anoxic zone (D) of Lake Bourget. 16S rRNA/16S rDNA ratios were presented for each group above the sequences proportions histograms.

Temporal distribution of active archaeal groups

Active archaeal communities retrieved from Lake Bourget were clearly dominated by *Thaumarchaeota* MGI and were stable through the whole sampling year. As no temporal variations were associated in active archaeal assemblage within this ecosystem, we decided to focus on Lake Pavin which showed community variation. This ecosystem presented temporal dynamics of the different active taxonomic groups in both the oxycline and the anoxic zone. In the oxycline (Figure 4.A), the MEG predominated active archaeal assemblage all year-round, except during March (27% of the reads) and November (11% of the reads).

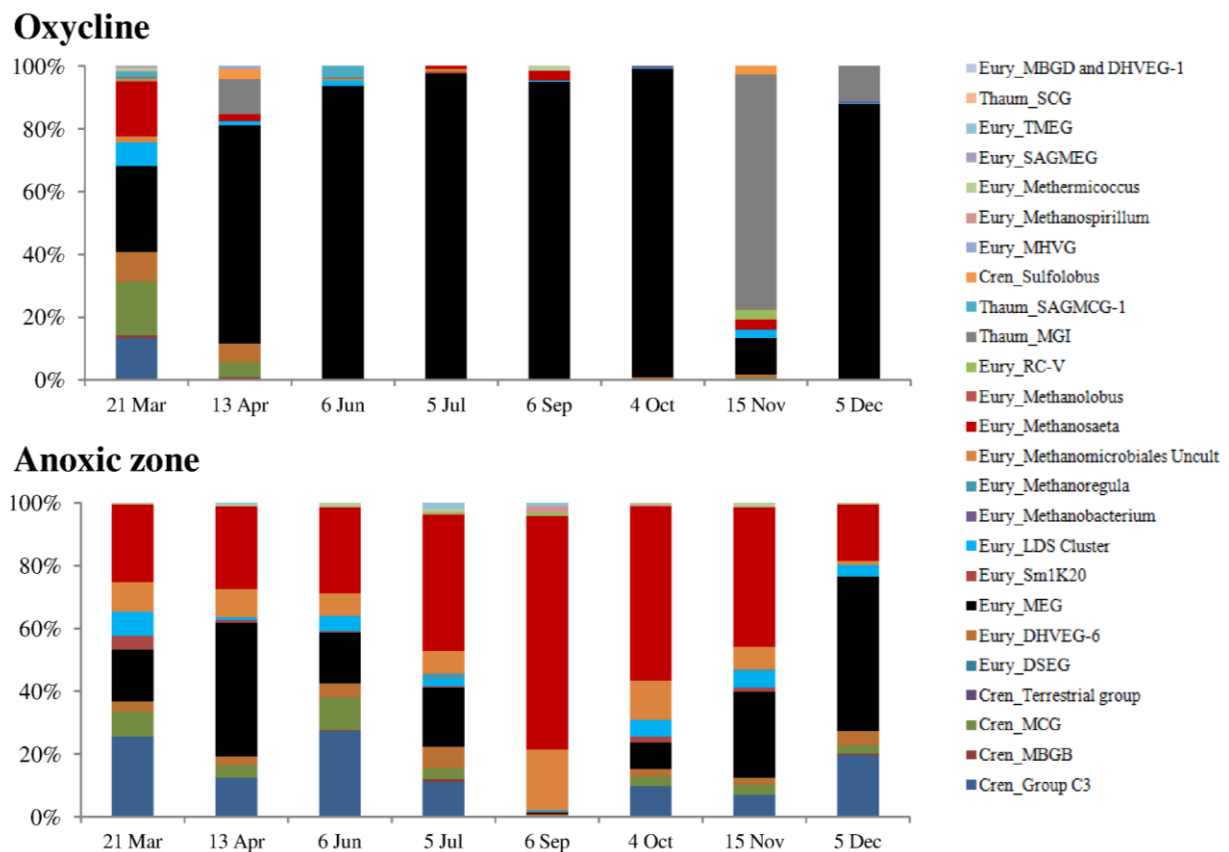


Figure 4. Temporal dynamics of active archaeal groups evaluated on the basis of the number of reads in Lake Pavin. Dynamics of active archaeal groups was followed in the oxycline (A) and the anoxic zone (B).

Thaumarchaeota MGI was more active during November and December than other periods. In the anoxic zone of Lake Pavin (Figure 4.B), *Methanoseta* were active all year round and represented the most active archaeal group during September (reaching 74% of the

reads), as *Methanomicrobiales* uncultured reads counts (about 19% of the reads), suggesting a dominance of active methanogenic lineages during this period. On the other hand, other groups like MEG or crenarchaeal Group C3 were active all year-round and MGI was never retrieved in this zone.

Statistical analyses conducted through the forward RDA indicated that among parameters recorded, ammonia, temperature, phosphate and oxygen explained 53.3, 23.4, 14.3 and 6.6% of archaeal communities' structure respectively ($P < 0.05$). This showed that oxygen had low influence on archaeal community's activity and that nitrogen compounds seemed to be the most structuring chemical parameter. In the same way, the more contrasted variation in oxygen concentrations recorded in Lake Bourget seemed not having an impact on the archaeal community since their composition was stable through the time. The RDA plot showed a clear difference between the activity of major taxonomic groups associated with Lake Bourget and those retrieved in Lake Pavin (Figure 5). Among significant parameters recorded with the forward RDA, temperature and oxygen were mostly linked to MGI communities and to Lake Bourget, while ammonia and phosphate were linked to different euryarchaeal lineages in Lake Pavin.

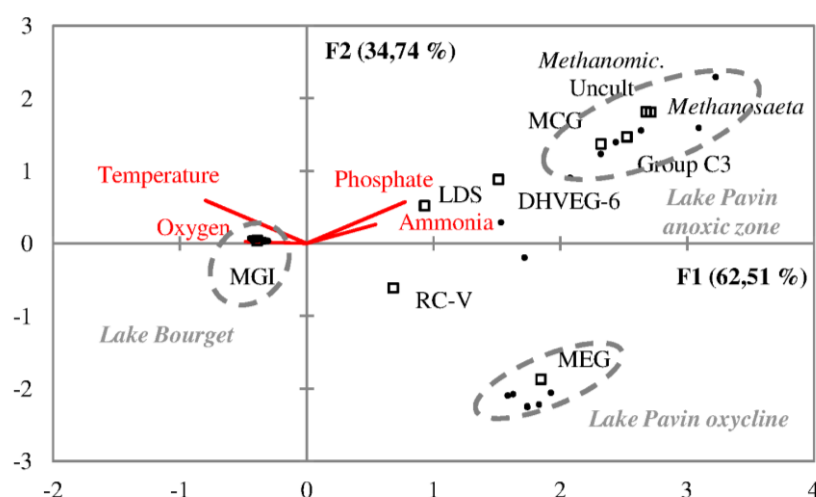


Figure 5. Ordination diagram from RDA of 16S rRNA major taxonomic groups compared with environmental data (temperature, oxygen, phosphate, ammonia).

Active ammonia oxidizing communities' dynamics

Thaumarchaeota MGI activity was not very important in the oxycline of Lake Pavin while active AOB were retrieved all year round (Figure 6.A).

In the oxycline zone of Lake Bourget, few archaeal and bacterial *amoA* transcripts were detected from March to October (less than 10 copies of transcripts.mL⁻¹). However an increase of both archaeal and bacterial *amoA* transcripts were observed during the winter period (Figure 6.B). In the anoxic zone, no AOA or AOB *amoA* transcripts were detected (data not shown).

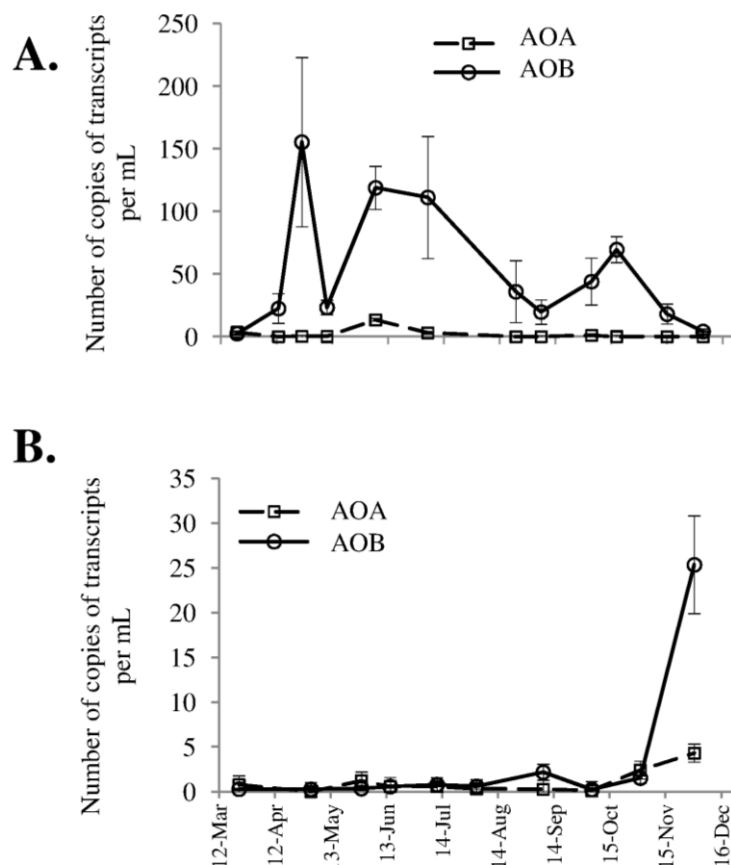


Figure 6. Archaeal and bacterial *amoA* transcripts dynamics in the oxycline zone of Lake Pavin (A) and Bourget (B).

Discussion

Archaea are key components of aquatic environments carrying out fundamental processes that influence ecosystem function and productivity (Pernthaler, 2005). Moreover, understanding the evolutionary or ecological parameters that shape freshwater archaeal diversity patterns in the water column is of major importance for ecological predictions and lacustrine management strategies. In this work, we focused on the meromictic Lake Pavin, usually considered as a stable ecosystem in deep layers with water masses well delineated and associated to particular prokaryotic communities (Lehours et al., 2005). Then we compared to a deep ecosystem, Lake Bourget which deep hypolimnion is considered as dynamic, since environmental parameters as O₂ and nutrients concentrations vary largely according to seasons (anoxic in early winter, mixing in February followed by progressive hypoxic conditions). Thus, the study of those two contrasted lakes in different water column zones, the oxycline and the anoxic zone, provided an ideal scheme to investigate the diversity and activity of archaeal communities.

An unexpected difference between archaeal compositions inferred from 16S rDNA and 16S rRNA studies highlighted specific active communities

This study clearly highlighted the importance of studying active communities. In this work, we demonstrated that the most abundant archaeal groups retrieved were not necessarily the most active, suggesting that ecological studies based only on a genomic approach might not be the reflect on microorganisms activity and ecosystems functioning. For example, the LDS and MEG groups were abundant in the anoxic zone of Lake Pavin and Bourget respectively while their activity inferred by the number of 16SrRNA sequences was very low. Since the growth rate was correlated with the number of ribosomes per cell (Campbell *et al.*, 2009;

Fegatella *et al.*, 1998), the ratio 16SrRNA/16SrDNA computed in this study allowed us to highlight which archaeal groups were the most active in the different zones studied.

Clearly, both ecosystems can be differentiated by their active communities dominated by *Thaumarchaeota* in Lake Bourget and *Euryarchaeota* in Lake Pavin. Interestingly, in the oxycline of Lake Pavin, the MEG was very active and clearly dominated the active archaeal assemblage through the whole year, suggesting that this poor-characterized group could play a key functional role in aquatic ecosystems. This group was usually retrieved in deep subsurface (Hirayama *et al.*, 2007) but also included clones from terrestrial soil and marine sediments (Takai *et al.*, 2001b).

In the anoxic zone of Lake Pavin, characterized methanogenic lineages present (*Methanosaeta* and *Methanomicrobiales*) were not the most abundant, however, those groups were the most active (based on the number of sequences and rRNA/rDNA ratios) confirming previous work suggesting that methanogenesis is the central process performed by *Archaea* in the monimolimnion of Lake Pavin (Borrel *et al.*, 2011). However, our study highlighted new archaeal groups in this ecosystem. Thus, the group C3 belonging to the *Crenarchaeota* were particularly active. This uncultured crenarchaeal group is often found in low-temperature terrestrial and marine habitats (DeLong and Pace, 2001) but also in deep marine sediments (Wang *et al.*, 2010) and saline lakes (Comeau *et al.*, 2012; Schneider *et al.*, 2013). Some other groups presented a slight mean seasonal activity as LDS or MCG. The LDS cluster was first identified in Lake Dagow sediment (Glissman *et al.*, 2004) and further identified in rivers (Galand *et al.*, 2006; Herfort *et al.*, 2009) where they could account for a large proportion of the archaeal cell counts (Herfort *et al.*, 2009). On the other hand, the MCG was one of the predominant archaeal groups obtained from marine deep subsurface sediments, but could also be retrieved in terrestrial, marine, hot and cold, surface and subsurface environments (Teske, 2006). A recent study suggested that this group could have a role in the carbon cycle (Biddle

350 *et al.*, 2006) but also in protein remineralization in anoxic marine sediments (Lloyd *et al.*, 2013).

Finally, the ratio 16SrRNA/16SrDNA assessed the great activity of the *Thaumarchaeota* MGI in Lake Bourget, while it was not active in Lake Pavin whatever the sampled zone.

***Thaumarchaeota* were not linked to nitrification process**

355 In Lake Bourget, active archaeal assemblage was clearly dominated by *Thaumarchaeota* Marine Group I, all year-round and at the two depths considered in oxic and anoxic conditions. Previous studies conducted in freshwater ecosystems demonstrated that the
360 oxycline zone was a favourable layer for archaeal ammonia oxidizers (Lliros *et al.*, 2010; Pouliot *et al.*, 2009). However, this parameter explained only a low variance in the archaeal community composition. The active MGI retrieved in Lake Bourget were mainly explained by low contents in ammonia and higher temperatures than those measured in Lake Pavin.

A recent study focused on archaeal activity in a saline meromictic lake assessing the
365 predominance of active MGI and AOA phylotypes in sub- and anoxic layers (La Cono *et al.*, 2013), as demonstrated in Lake Bourget. Thus, by targeting AOA in Lake Bourget, we hypothesised that those microorganisms could be active in this lake. Nevertheless, our study showed that archaeal and bacterial *amoA* transcripts abundance were in average low during the whole year in Lake Bourget. These results suggested therefore that among lakes the
370 oxycline was not obligatory a hotspot for archaeal ammonia oxidation and that MGI present and active could use another metabolic pathway as energy source. Indeed, it has been suggested that those microorganisms presented a large metabolic plasticity, from autotrophy to possibly mixotrophic lifestyles (Herndl *et al.*, 2005; Konneke *et al.*, 2005). In the oxycline and the anoxic zone of this lake, the aerobic nitrifiers would likely be replaced by anaerobic

ammonium oxidation (anammox) bacteria, such as those detected in deep anaerobic waters of the Black Sea (Kuypers *et al.*, 2003). Nevertheless, further molecular analysis targeting nitrifying, denitrifying and anammox *Bacteria* were needed to identify the community of nitrogen cycling microorganisms in lakes.

Moreover, AOA and AOB might compete for ammonia oxidation and it has been demonstrated that low-ammonia concentrations were favourable for AOA activity (Hatzenpichler, 2012; Schleper and Nicol, 2010). This was congruent with the absence of activity for MGI in Lake Pavin where ammonia concentrations were higher. In this lake ammonia oxidation was almost performed by AOB that were active during the whole year, and while the deeper monimolimnion was rich in sulfide, it was previously shown that AOA, or at least some ecotypes are likely to be tolerant to sulfide and able to oxidize ammonia in its presence, suggesting that in this zone, the oxygen depletion coupled with high ammonia levels might be the most important factors that prevent their activity. Though, several environmental parameters need to be considered to explain AOA distribution and activity (Erguder *et al.*, 2009; Hatzenpichler, 2012).

The dynamics of *Archaea* in the monimolimnion was greater than expected

The permanent anoxic zone of Lake Pavin was characterized by an important archaeal 16SrRNA activity and certainly evidenced that this permanent depletion in oxygen favoured particularly *Euryarchaeota* that were diverse compared to the anoxic zone of the Lake Bourget dominated by *Thaumarchaeota*. In addition, the monimolimnion of meromictic lakes should not be considered as particularly stable in their microbial composition as evidenced in this study and as shown by long term constancy. A previous study conducted punctually in 2002 in Lake Pavin on the same transition oxycline to anoxic zone, recorded through *in situ* hybridization the presence of approximately 6.10^6 archaeal cells.mL⁻¹ (Lehours *et al.*, 2005).

400 This discrepancy could be due to methodological bias, however, ten year separated the two studies, and total prokaryotes cytometric counts (data not shown) revealed the decrease of abundance of prokaryotic groups present in this lake (Colombet *et al.*, 2009), that could be congruent with the decrease in archaeal cell counts observed in our study.

Although, unexpected temporal changes were recorded in the monimolimnion of Lake
405 Pavin. Some hypotheses could explain these temporal variations as sedimentation processes. Indeed, the inactive *Thaumarchaeota* MGI sequences represented an important part of the present archaeal assemblages (22.8% of the reads in the 16S rDNA dataset). However, the sedimentation flux could not explain temporal changes observed for the active archaeal fraction and environmental parameters recorded during this work suggested an important
410 variability of phosphate, ammonia or even nitrate in the anoxic zone of Lake Pavin, which could impact active archaeal rearrangement in this zone.

Nevertheless, in this study we focused only on bottom-up controls that could occur in this ecosystem while predation and viral lysis can control the population diversity in aquatic ecosystem (Pernthaler, 2005). Indeed, viral communities were retrieved in the permanently
415 anoxic monimolimnion of Lake Pavin and could thus significantly contribute to the regulation of prokaryotic communities (Colombet *et al.*, 2009). In this case, we could imagine that archaeal temporal rearrangements that occurred in Lake Pavin could be the result of viral lysis of specific taxonomic groups.

Indeed, it has been suggested that the ecology of the deepest waters of Lake Pavin was
420 essentially driven by the dark viral loop (DOM-prokaryotes-viruses) processes, which could sequester organic matters and nutrients for a long-live turnover (Colombet *et al.*, 2009). Archaeal communities' rearrangements could be then explained by the "killing the winner" hypothesis where winners refers not necessarily to the most abundant but to the most active prokaryotic populations (Winter *et al.*, 2010).

425 Our work revealed clearly different patterns of active archaeal assemblages in two
contrasted freshwater ecosystems, and point out that previous studies and observations should
not be generalized, as archaeal distribution and activity could be very different. We note that
taxa present in small proportions but active might play key functional role in freshwater
ecosystems and that high-throughput studies coupled to single-cell approaches will allow
430 gaining insights into their ecological role.

Acknowledgments

Physicochemical data were analyzed by the observatory on peri-alpine lakes (Database
SOERE GLACPE INRA, UMR CARTELE, Thonon les Bains, and data CISALB). We thank
435 G. Paolini, P. Perney, L. Lainé and L. Jaccas for their technical contributions to sampling,
analysis of physical and chemical parameters in Lake Bourget, and contribution to water
samples' preparation. We thank S. Palesse and J. Colombet for their support on the field in
Lake Pavin, A. Vellet, A. Moné and I. Louati for physical and chemical analyses, and
contribution to molecular analyses.

440

445

References

- Auguet JC, Nomokonova N, Camarero L, Casamayor EO (2011). Seasonal changes of
450 freshwater ammonia-oxidizing archaeal assemblages and nitrogen species in oligotrophic
alpine lakes. *Appl Environ Microbiol* **77**: 1937-45.
- Biddle JF, Lipp JS, Lever MA, Lloyd KG, Sorensen KB, Anderson R *et al* (2006).
455 Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl
Acad Sci USA* **103**: 3846-51.
- Borcard D, Legendre P, Drapeau P (1992). Partialling out the spatial component of ecological
variation. *Ecology* **73**: 1045-1055.
- 460 Bosshard PP, Santini Y, Gruter D, Stettler R, Bachofen R (2000). Bacterial diversity and
community composition in the chemocline of the meromictic alpine Lake Cadagno as
revealed by 16S rDNA analysis. *FEMS Microbiol Ecol* **31**: 173-182.
- Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011). Activity of abundant and rare
465 *bacteria* in a coastal ocean. *Proc Natl Acad Sci USA* **108**: 12776-81.
- Campbell BJ, Yu L, Straza TRA, Kirchman DL (2009). Temporal changes in bacterial rRNA
and rRNA genes in Delaware (USA) coastal waters. *Aqua Microbial Ecol* **57**: 123-135.
- 470 Colombet J, Charpin M, Robin A, Portelli C, Amblard C, Cauchie HM *et al* (2009). Seasonal
depth-related gradients in virioplankton: standing stock and relationships with microbial
communities in Lake Pavin (France). *Microb Ecol* **58**: 728-36.
- Comeau AM, Harding T, Galand PE, Vincent WF, Lovejoy C (2012). Vertical distribution of
475 microbial communities in a perennially stratified Arctic lake with saline, anoxic bottom
waters. *Sci Rep* **2**: 604.
- DeLong EF, Pace NR (2001). Environmental diversity of *bacteria* and *archaea*. *Syst Biol* **50**:
470-8.
- 480 Erguder TH, Boon N, Wittebolle L, Marzorati M, Verstraete W (2009). Environmental factors
shaping the ecological niches of ammonia-oxidizing *archaea*. *FEMS Microbiol Rev* **33**: 855-
69.
- 485 Fegatella F, Lim J, Kjelleberg S, Cavicchioli R (1998). Implications of rRNA operon copy
number and ribosome content in the marine oligotrophic ultramicrobacterium *Sphingomonas*
sp. strain RB2256. *Appl Environ Microbiol* **64**: 4433-8.
- 490 Galand PE, Bourrain M, De Maistre E, Catala P, Desdevises Y, Elifantz H *et al* (2012).
Phylogenetic and functional diversity of *Bacteria* and *Archaea* in a unique stratified lagoon,
the Clipperton atoll (N Pacific). *FEMS Microbiol Ecol*.
- Galand PE, Lovejoy C, Vincent WF (2006). Remarkably diverse and contrasting archaeal
495 communities in a large arctic river and the coastal Arctic Ocean. *Aquatic Microbial Ecology*
44: 115-126.

- Glissman K, Chin KJ, Casper P, Conrad R (2004). Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature. *Microb Ecol* **48**: 389-99.
- 500 Glockner FO, Fuchs BM, Amann R (1999). Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Appl Environ Microbiol* **65**: 3721-6.
- 505 Grossart HP, Frindte K, Dziallas C, Eckert W, Tang KW (2011). Microbial methane production in oxygenated water column of an oligotrophic lake. *Proc Natl Acad Sci U S A* **108**: 19657-61.
- 510 Großkopf R, Stubner S, Liesack W (1998). Novel Euryarchaeotal Lineages Detected on Rice Roots and in the Anoxic Bulk Soil of Flooded Rice Microcosms. *Appl Environ Microbiol* **64**: 4983-4989.
- 515 Hatzenpichler R (2012). Diversity, physiology and niche differentiation of ammonia-oxidizing archaea. *Appl Environ Microbiol*.
- Hatzenpichler R, Lebedeva EV, Spieck E, Stoecker K, Richter A, Daims H *et al* (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci USA* **105**: 2134-9.
- 520 Herfort L, Kim JH, Coolen MJL, Abbas B, Schouten S, Herndl GJ *et al* (2009). Diversity of *Archaea* and detection of crenarchaeotal *amoA* genes in the river Rhine and Têt. *Aquatic Microbial Ecology* **55**: 189-201.
- 525 Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A *et al* (2005). Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303-9.
- 530 Hirayama H, Sunamura M, Takai K, Nunoura T, Noguchi T, Oida H *et al* (2007). Culture-dependent and -independent characterization of microbial communities associated with a shallow submarine hydrothermal system occurring within a coral reef off Taketomi Island, Japan. *Appl Environ Microbiol* **73**: 7642-56.
- 535 Hugoni M, Etien S, Bourges A, Lepere C, Domaizon I, Mallet C *et al* (2013a). Dynamics of ammonia-oxidizing *Archaea* and *Bacteria* in contrasted freshwater ecosystems. *Res Microbiol* **164**: 360-70.
- 540 Hugoni M, Taib N, Debroas D, Domaizon I, Jouan Dufournel I, Bronner G *et al* (2013b). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci USA* **110**: 6004-9.
- Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K *et al* (2003). Microbial communities associated with geological horizons in coastal subseafloor sediments from the sea of okhotsk. *Appl Environ Microbiol* **69**: 7224-35.
- 545 Jones SE, Lennon JT (2010). Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci USA* **107**: 5881-6.

- Jurgens G (2000). Molecular phylogeny of *Archaea* in boreal forest soil, freshwater and temperate estuarine sediment. *Thèse*.
- 550 Keough BP, Schmidt TM, Hicks RE (2003). Archaeal nucleic acids in picoplankton from great lakes on three continents. *Microb Ecol* **46**: 238-48.
- Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA (2005).
555 Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543-6.
- Kowalchuk GA, Stephen JR (2001). Ammonia-oxidizing bacteria: a model for molecular microbial ecology. *Annu Rev Microbiol* **55**: 485-529.
- 560 Kuypers MM, Sliekers AO, Lavik G, Schmid M, Jorgensen BB, Kuenen JG *et al* (2003). Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. *Nature* **422**: 608-11.
- La Cono V, La Spada G, Arcadi E, Placenti F, Smedile F, Ruggeri G *et al* (2013). Partaking of *Archaea* to biogeochemical cycling in oxygen-deficient zones of meromictic saline Lake Faro (Messina, Italy). *Environ Microbiol* **15**: 1717-33.
- 565 Lami R, Ghiglione JF, Desdevises JF, West NJ, Lebaron P (2009). Annual patterns of presence and activity of marine bacteria monitored by 16S rDNA-16SrRNA fingerprints in the coastal NW Mediterranean Sea. *Aquat Microb Ecol* **54**: 199-210.
- 570 Lehours AC, Bardot C, Thenot A, Debroas D, Fonty G (2005). Anaerobic microbial communities in Lake Pavin, a unique meromictic lake in France. *Appl Environ Microbiol* **71**: 7389-400.
- 575 Lliros M, Casamayor EO, Borrego C (2008). High archaeal richness in the water column of a freshwater sulfurous karstic lake along an interannual study. *FEMS Microbiol Ecol* **66**: 331-342.
- Lliros M, Gich F, Plasencia A, Auguet JC, Darchambeau F, Casamayor EO *et al* (2010).
580 Vertical distribution of ammonia-oxidizing *crenarchaeota* and methanogens in the epipelagic waters of Lake Kivu (Rwanda-Democratic Republic of the Congo). *Appl Environ Microbiol* **76**: 6853-63.
- Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD *et al* (2013).
585 Predominant *archaea* in marine sediments degrade detrital proteins. *Nature* **496**: 215-8.
- Lorenzen CJ (1967). Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnology and Oceanography* **12**: 343-346.
- 590 Pernthaler J (2005). Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3**: 537-46.
- Pernthaler J, Glockner FO, Unterholzner S, Alfreider A, Psenner R, Amann R (1998).
595 Seasonal community and population dynamics of pelagic *bacteria* and *archaea* in a high mountain lake. *Appl Environ Microbiol* **64**: 4299-306.

Pouliot J, Galand PE, Lovejoy C, Vincent WF (2009). Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two meromictic High Arctic lakes. *Environ Microbiol* **11**: 687-99.

Schleper C, Nicol GW (2010). Ammonia-oxidising archaea--physiology, ecology and evolution. *Adv Microb Physiol* **57**: 1-41.

Schneider D, Arp G, Reimer A, Reitner J, Daniel R (2013). Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the kiritimati atoll, central pacific. *PLoS One* **8**: e66662.

Strickland J, Parsons T (1968). *A practical handbook of sea water analysis*.

Taib N, Mangot JF, Domaizon I, Bronner G, Debroas D (2013). Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS One* **8**: e58950.

Takai K, Moser DP, DeFlaun M, Onstott TC, Fredrickson JK (2001). Archaeal diversity in waters from deep South African gold mines. *Appl Environ Microbiol* **67**: 5750-60.

Teske A (2006). Microbial communities of deep marine subsurface sediments: molecular and cultivation surveys. *Geomicrobiology Journal* **23**: 357-368.

Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic *crenarchaeota* in nitrogen cycling. *Environ Microbiol* **7**: 1985-95.

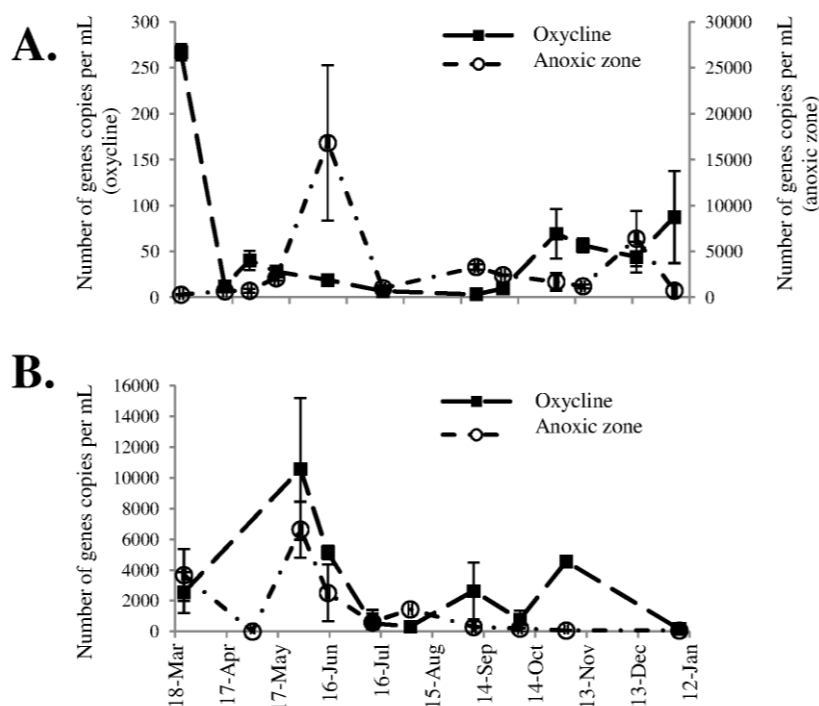
Viollier E, Michard G, Jezequel D, Pepe M, Sarazin G (1997). Geochemical study of a crater lake: Lake Pavin, Puy de Dôme, France. Constraints afforded by the particular matter distribution in the element cycling within the lake. *Chemical Geology* **142**: 225-241.

Vissers EW, Anselmetti FS, Bodelier PL, Muyzer G, Schleper C, Tourna M *et al* (2013). Temporal and spatial coexistence of archaeal and bacterial amoA genes and gene transcripts in Lake Lucerne. *Archaea* **2013**: 289478.

Wang P, Li T, Hu A, Wei Y, Guo W, Jiao N *et al* (2010). Community structure of *archaea* from deep-sea sediments of the South China Sea. *Microb Ecol* **60**: 796-806.

Winter C, Bouvier T, Weinbauer MG, Thingstad TF (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiol Mol Biol Rev* **74**: 42-57.

Supplementary data



Supplementary Figure 1. Abundances of 16S rDNA genes per mL in Lake Pavin (A) and Lake Bourget (B). Transcripts copy numbers were quantified by qPCR over the course of a whole year among two sampling depths, 45 m (oxycline) and 80 m (anoxic zone) for Lake Pavin and 130 m (oxycline) and 140 m (anoxic zone) for Lake Bourget.

Supplementary Table 1. Primers used for pyrosequencing and RT-qPCR in this study.

Application	Primer	Primer sequence 5' – 3'	Annealing temperature	Targeted gene	Reference
Pyrosequencing	Arch519F	CAGCCGCCGCGGTAA	57°C	Archaeal 16S rRNA	Herfort <i>et al.</i> , 2009
	Arch915R	GTGCTCCCCGCCAATTCCT	57°C	Archaeal 16S rRNA	Casamayor <i>et al.</i> , 2002
qPCR	CrenAmoAModF	TGGCTAAGACGMTGTA	52°C	Thaumarchaeal <i>amoA</i>	Mincer <i>et al.</i> , 2007
	CrenAmoAModR	AAGCGGCCATCCATCTGTA	52°C	Thaumarchaeal <i>amoA</i>	Mincer <i>et al.</i> , 2007
	amoA-1F	GGGGTTTCTACTGGTGGT	56°C	β-proteobacterial <i>amoA</i>	Rotthauwe <i>et al.</i> , 1997
	AmoA-RNEW	CCCCTCBGSAAAVCCTTCTTC	56°C	β-proteobacterial <i>amoA</i>	Rotthauwe <i>et al.</i> , 1997

Supplementary Table 2. Quality checked (QC) and archaeal affiliated read counts obtained for each sample from the oxycline and anoxic zone in Lake Pavin and Lake Bourget.

Sampling Site	Sampling Date	QC sequences		Archaeal sequences	
		16S rDNA	16S rRNA	16S rDNA	16S rRNA
Lake Pavin oxycline	21-Mar-11	11741	3794	9660	2117
	13-Apr-11	1893	10637	58	432
	26-Apr-11	1514	9782	32	746
	10-May-11	1308	ND	3	ND
	6-Jun-11	1502	5439	13	300
	5-Jul-11	2571	6505	95	2846
	23-Aug-11	3485	ND	223	ND
	6-Sep-11	5134	7971	3399	1895
	4-Oct-11	8327	4791	3061	1464
	18-Oct-11	2397	3355	976	402
	15-Nov-11	3772	12755	231	1401
	5-Dec-11	11281	17117	5445	9618
Lake Pavin anoxic zone	21-Mar-11	2630	777	2426	571
	13-Apr-11	7132	1248	6027	499
	10-May-11	5357	708	4916	568
	6-Jun-11	5339	607	5020	395
	5-Jul-11	2097	1154	1847	1020
	23-Aug-11	641	1126	602	903
	6-Sep-11	1267	1402	1140	1208
	4-Oct-11	ND	1344	ND	1141
	15-Nov-11	ND	1263	ND	1116
	5-Dec-11	ND	566	ND	510
Lake Bourget oxycline	23-Mar-11	18473	ND	12898	ND
	30-May-11	26	4397	12	4229
	15-Jun-11	3018	3159	1983	3039
	12-Jul-11	2555	1343	2422	1268
	2-Aug-11	2699	2425	2664	2331
	8-Sep-11	15275	ND	14888	ND
	5-Oct-11	2532	ND	2451	ND
	15-Nov-11	3097	1445	2841	1375
	15-Dec-11	4195	4441	4001	4184
Lake Bourget anoxic zone	23-Mar-11	3275	15418	3188	14558
	30-May-11	4944	2115	26	1929
	15-Jun-11	3	1326	3	1247
	12-Jul-11	8	22086	4	19981
	2-Aug-11	2	4	0	4
	8-Sep-11	11106	17245	784	15571
	5-Oct-11	5260	10984	2275	7244
	15-Nov-11	3484	1371	2310	513
	15-Dec-11	ND	1853	ND	1796

Supplementary Table 3. Environmental parameters (pH, temperature and oxygen, phosphate, nitrate and ammonia concentrations) associated with the oxycline and the anoxic zone of each lake. Average values were presented, with Min-Max values retrieved during the sampling year (9 sampling points in Lake Pavin and 12 in Lake Bourget) and the CV. ND, not determined.

	pH	Temp. °C	Oxy. mg.L ⁻¹	P-PO ₄ ³⁻ mgP.L ⁻¹	N-NO ₃ ⁻ mgN.L ⁻¹	N-NH ₄ ⁺ mgN.L ⁻¹
Lake Pavin oxycline						
Average values	6.81	4.23	1.73	0.05	0.11	0.42
Min-Max	6.23-7.14	4.20-4.50	1.0-2.89	0-0.09	0-0.45	0.04-0.85
CV (%)	5	2	38	238	138	70
Lake Pavin anoxic zone						
Average values	6.22	4.70	0.43	3.3	0.16	31.21
Min-Max	5.87-7.09	4.20-5.10	0.41-0.55	0-6.10	0-0.87	5.29-98.46
CV (%)	6	9	19	58	165	115
Lake Bourget oxycline						
Average values	7.49	5.48	6.15	0.03	0.56	0
Min-Max	7.34-7.86	5.47-5.50	3.24-9.47	0.01-0.04	0.45-0.64	0-0.008
CV (%)	2	0	35	37	14	64
Lake Bourget anoxic zone						
Average values	7.34	5.50	2.82	0.04	0.55	0.09
Min-Max	7.18-7.85	5.43-5.52	0.06-9.37	0.01-0.06	0.47-0.75	0.001-0.24
CV (%)	3	1	122	49	10	135

CHAPITRE 2 : Impact de gradients environnementaux sur la diversité spécifique et fonctionnelle des *Archaea*

5.1 Contexte et objectifs

Dans le chapitre précédent, nous avons pu mettre en évidence des différences notables dans la structuration des communautés actives d'*Archaea* principalement sur la base de l'étude du marqueur phylogénétique ADN_r16S. Par ailleurs, ces microorganismes, du fait de leur abondance et leur ubiquité dans les écosystèmes aquatiques sont des acteurs des différents cycles biogéochimiques. La bibliographie fait état de l'implication des *Thaumarchaeota* du MGI dans le cycle de l'azote et plus particulièrement l'oxydation de l'ammonium, cependant les facteurs environnementaux qui régulent les communautés d'AOA sont encore mal connus. La structuration des communautés de *Thaumarchaeota* et plus particulièrement celle du MGI, connu pour oxyder l'ammonium en milieu marin, avec une abondance et une activité plus importante en période hivernale, nous a conduits à nous interroger sur les facteurs pouvant impacter ce trait métabolique. En outre, peu de données sont à l'heure actuelle disponible quant à leur activité ou leur contribution relative à ce processus par rapport aux bactéries oxydant l'ammonium (AOB). Dans ce chapitre, deux études différentes ont été conduites, l'une sur des milieux lacustres (Article 3), l'autre dans un estuaire (Article 4), et avaient pour objectifs d'évaluer l'impact de gradients environnementaux (anthropiques et salins) sur la structuration et l'activité des communautés d'AOA, en comparaison avec celles des AOB.

L'étude de milieux lacustres contrastés (Article 3) nous a permis d'attester la présence des AOA et des AOB dans des milieux d'eau douce de niveau trophique différent. Ceci nous a permis d'évaluer quels étaient les paramètres environnementaux susceptibles d'impacter l'établissement des communautés d'AOA, mais également d'évaluer leurs variations temporelles au cours de deux périodes annuelles contrastées : la période hivernale et la période estivale.

D'autre part, pour étudier l'influence possible de la salinité sur la structure des *Archaea* actives, nous avons porté notre choix sur un écosystème estuarien (Article 4) situé à l'interface terre-mer, caractérisé par des gradients physico-chimiques drastiques (tant en terme de salinité, de turbidité que de concentrations en nutriments) susceptibles d'impacter les communautés d'*Archaea* mais aussi plus particulièrement les AOA. Dans le cadre de ce travail, nous nous sommes focalisés sur la structuration des communautés archéennes actives (basé sur l'analyse des ARN_r 16S) en prenant en compte leur dynamique temporelle dans 3 stations d'échantillonnage de salinité différente. Ainsi, nous avons pu tester la présence

d'écotypes au sein des populations d'*Archaea* actives et porter un accent particulier sur la dynamique temporelle des AOA et AOB actives. Nous avons par ailleurs étudié la dynamique des gènes actifs impliqués dans la dégradation de l'urée, processus récemment mis en évidence chez les AOA et décrit comme pouvant apporter une source d'ammonium et de carbone aux AOA (Alonso-Saez *et al.*, 2012; Konstantinidis *et al.*, 2009; Yakimov *et al.*, 2011).

5.2 Article 3 : Dynamics of ammonia-oxidizing *Archaea* and *Bacteria* in contrasted freshwater ecosystems

Mylène Hugoni, Sandrine Etien, Antoine Bourges, Cécile Lepère, Isabelle Domaizon,
Clarisse Mallet, Gisèle Bronner, Didier Debroas, Isabelle Mary

Research in Microbiology

2013

Dynamics of ammonia-oxidizing *Archaea* and *Bacteria* in contrasted freshwater ecosystems

Mylène Hugoni^{a,b}, Sandrine Etien^{a,b}, Antoine Bourges^{a,b}, Cécile Lepère^{a,b}, Isabelle Domaizon^c,
Clarisse Mallet^{a,b}, Gisèle Bronner^{a,b}, Didier Debroas^{a,b,*}, Isabelle Mary^{a,b}

^a Clermont Université, Université Blaise Pascal, Laboratoire “Microorganismes: Génome et Environnement”, BP 10448, F-63000 Clermont-Ferrand, France

^b CNRS, UMR 6023, LMGE, F-63171 Aubière, France

^c INRA, UMR CARTEL, F-74200 Thonon-Les-Bains, France

Received 10 September 2012; accepted 21 January 2013

Available online 8 February 2013

Abstract

Thaumarchaeota have been recognized as the main drivers of aerobic ammonia oxidation in many ecosystems. However, little is known about the role of ammonia-oxidizing *Archaea* (AOA) and *Bacteria* (AOB) in lacustrine ecosystems. In this study, the photic zone of three contrasted freshwater ecosystems located in France was sampled during two periods: winter homothermy (H) and summer thermal stratification (TS), to investigate the distribution of planktonic AOA and AOB. We showed that AOB were predominant in nutrient-rich ecosystems, whereas AOA dominated when ammonia concentrations were the lowest and during winter, which could provide a favorable environment for their growth. Moreover, analyses of archaeal libraries revealed the ubiquity of the thaumarchaeal I.1a clade associated with higher diversity of AOA in the most nutrient-poor lake. More generally, this work assesses the presence of AOA in lakes, but also highlights the existence of clades typically associated with lacustrine and hot spring ecosystems and specific ecological niches occupied by these microorganisms.

© 2013 Published by Elsevier Masson SAS on behalf of Institut Pasteur.

Keywords: Ammonia oxidation; *amoA*; *Thaumarchaeota*; *Bacteria*; Lakes; 16S rRNA

1. Introduction

Aerobic nitrification of ammonia to nitrite and nitrate is an essential process in the global nitrogen cycle. Historically, aerobic ammonia oxidation, the first rate-limiting step of nitrification, was thought to be performed exclusively by ammonia-oxidizing *Bacteria* (AOB), particularly *Betaproteobacteria*, which are widely distributed in aquatic and terrestrial ecosystems (Kowalchuk and Stephen, 2001). However, metagenomic studies (Treusch et al., 2005; Venter et al., 2004) suggested that non-thermophilic *Crenarchaeota*, now classified

as *Thaumarchaeota* (Brochier-Armanet et al., 2008), may also play a role in ammonia oxidation. This was confirmed by detection in numerous microbial assemblages of the archaeal *amoA* gene encoding a subunit of the key enzyme involved in ammonia oxidation, ammonia monooxygenase (Amo) (Francis et al., 2007; Schleper et al., 2005; Venter et al., 2004). Indeed, several studies have revealed the existence of *amoA* genes in members of the thaumarchaeal I.1b (Treusch et al., 2005) and I.1a groups (also named Marine Group I, (Hallam et al., 2006; Könneke et al., 2005). *Archaea* have increasingly become recognized as abundant in marine (DeLong, 1992; Karner et al., 2001) and terrestrial environments (Leininger et al., 2006), and archaeal *amoA* abundance (ammonia-oxidizing *Archaea*: AOA) can be 1–2 orders of magnitude higher than bacterial *amoA* (Leininger et al., 2006). Recent studies suggest that the relative importance of AOA and AOB may be closely related to environmental factors such as oxygen concentration

* Corresponding author. LMGE, Laboratoire Microorganismes: Génome et Environnement, UMR CNRS 6023, Université Blaise Pascal (Clermont-Ferrand II), 24 avenue des Landais, BP 80026, Aubière 63171, France. Tel.: +33 4 73 40 78 37; fax: +33 4 73 40 76 70.

E-mail address: didier.debroas@univ-bpclermont.fr (D. Debroas).

(Lliros et al., 2010; Pouliot et al., 2009), ammonia levels (Hatzenpichler et al., 2008; Park et al., 2006), salinity (Francis et al., 2005; Mosier and Francis, 2008), pH, nitrite, nitrate and phosphate concentrations (Erguder et al., 2009). Nevertheless, comparative ecological studies need to be conducted in various ecosystems to fully understand ammonia-oxidizing community distribution and dynamics and factors controlling these communities. While most published surveys have been conducted on marine and terrestrial ecosystems, few studies have explored the role of *Thaumarchaeota* in the nitrogen cycle in lacustrine ecosystems (Auguet et al., 2011; Lliros et al., 2010). Nevertheless, the presence of metabolically active *Archaea* is now well established in the photic zone of these ecosystems (Boucher et al., 2006; Jardillier et al., 2005) and high numbers of *Archaea* (up to 37% of prokaryotic cell counts), mainly *Thaumarchaeota* group I.1a, were also observed at the air–water interface (neuston) of oligotrophic high-altitude mountain lakes (Auguet and Casamayor, 2008). Moreover, a study conducted in three oligotrophic high mountain lakes showed temporal changes in AOA diversity in surface waters and the abundance of this group was linked to changes in ammonium and nitrite concentrations (Auguet et al., 2011). However, the relative contribution of AOA vs. AOB in ammonia oxidation remains unknown in lacustrine systems. In this study, we investigated the dynamics of AOA and AOB (i.e. *amoA* genes) in three lakes differing in their nutrient concentrations and their watershed, but which were all characterized by winter homothermy and thermal stratification periods.

2. Materials and methods

2.1. Study sites, sampling and chemical analyses

The three lakes investigated are located in France and their main characteristics are reported in Table 1. Lake Bourget is known as the largest French lake, while the Sep reservoir is an artificial lake with high anthropogenic input and water renewal and Lake Aydat is a nutrient-rich ecosystem. Samples were collected monthly in 2002 (Lake Aydat), 2006 (Sep reservoir)

and 2009 (Lake Bourget) for quantitative analysis of thaumarchaeal 16S rRNA, archaeal and betaproteobacterial *amoA* (except for March and April when the Sep reservoir and Lake Aydat were frozen and inaccessible). Archaeal 16S rRNA and *amoA* gene libraries were constructed from a sample collected during winter homothermy (from November to February) and summer thermal stratification (from May to September). Water samples were collected in the epilimnion (2 m) of the Sep reservoir and Lake Aydat using a Van Dorn bottle at a permanent station located at the deepest zone of the water column. The 0–20 m layer was sampled in Lake Bourget using an integrating water sampler to carry out sampling right through the water column (Pelletier and Orand, 1978). Water temperature and dissolved oxygen content were determined with a multiparameter probe (YSI GRANT 3800). Phosphorus (P-PO₄), nitrate (N-NO₃) and ammonium (N-NH₄) contents were analyzed using standard American Public Health Association (1992) methods. Chlorophyll-*a* (Chl_a) content was determined by spectrophotometry (Lorenzen, 1967; Strickland and Parsons, 1968).

2.2. DNA extraction

A subsample was pre-filtered through 5 µm pore-size polycarbonate filters (Millipore), collected on 0.2-µm pore-size (pressure <10 kPa) polycarbonate filters (Millipore) and stored at –80 °C until nucleic acid extraction. The extraction procedure used was previously described in Lefranc et al. (2005). Briefly, DNA was extracted by digesting the cells with lysozyme (final concentration: 2 µg µL^{–1}) and proteinase K (100 µg mL^{–1}). Nucleic acids were then separated with phenol/chloroform and precipitated with ethanol. DNA yield was quantified using a Nanodrop spectrophotometer (ND 1000 Spectrophotometer) and nucleic acid extracts were stored at –20 °C until analysis.

2.3. Quantitative PCR analysis

qPCR analyses were performed on a Mastercycler ep gradientS Realplex² Eppendorf thermocycler running Realplex software (version 1.5) using the primers and conditions listed in Table 2. Copy numbers of thaumarchaeal 16S rRNA, archaeal *amoA* and betaproteobacterial *amoA* in the environmental samples were determined in duplicate on the non-diluted sample and for two different sample dilutions (5-fold and 25-fold). The reaction mixture (25 µL) contained MESA GREEN qPCR MasterMix Plus for SYBR assay (1X, Eurogentec). This mix contained dNTPs, Meteor Taq DNA polymerase, MgCl₂ and SYBR Green. For a 25 µL reaction mix, 8 µg of BSA, 0.2 µM of primers and ultra-pure sterile water were added. All reactions were performed in 96-well qPCR plates (Eurogentec) with optical tape (Eurogentec). One µl of diluted (5-fold and 25-fold) or non-diluted environmental DNA was added to 24 µl of mix in each well. The accumulation of newly amplified double-stranded gene products was followed online as the increase in fluorescence due to binding of SYBR Green fluorescent dye. All qPCR reactions consisted

Table 1
Main characteristics of the epilimnion of the different lakes sampled.

	Lake Bourget	Sep reservoir	Lake Aydat
Coordinates	45° 44' N 5° 51' E	46° 01' N 3° 01' E	45° 39' N 3° 01' E
Altitude (m)	231	503	825
Area (km ²)	44.5	3.3	6.03
Maximum depth (m)	145.5	37.4	14.5
	H–TS	H–TS	H–TS
N-NH ₄ (mgN L ^{–1})	0.003–0.01	0.05–0.04	0.29–0.11
N-NO ₃ (mgN L ^{–1})	0.45–0.47	0.9–0.9	4.18–4.21
P-PO ₄ (mgP L ^{–1})	0.002–0.002	0.01–0.03	0.06–0.04
Temperature (°C)	10.01–15.68	13.9–21.7	10.97–18.81
Oxygen (mg L ^{–1})	10.64–10.25	8.37–8.60	10.47–10.45
Chl _a (µg L ^{–1})	6.93–5.12	9.9–11.6	10.6–6.29

Environmental parameters such N-NH₄, N-NO₃, P-PO₄, temperature, oxygen, and Chl_a were followed during the entire year. Average values found during winter homothermy (H) and summer thermal stratification (TS) are presented.

Table 2
PCR primers and annealing conditions used in this study.

Application	Primer	Primer sequence 5'–3'	Annealing temperature	Targeted group	Reference
Cloning	Ar4f	TCYGGTTGATCCTGCCRG	56 °C	Archaeal 16S rRNA	Jurgens et al., 2000
	Ar958r	YCC GGC GTT GAV TCC AAT T	56 °C	Archaeal 16S rRNA	Jurgens et al., 2000
	Arch-amoAF	STA ATG GTC TGG CTT AGA CG	53 °C	Thaumarchaeal <i>amoA</i>	Francis et al., 2005
	Arch-amoAR	GCG GCC ATC CAT CTG TAT GT	53 °C	Thaumarchaeal <i>amoA</i>	Francis et al., 2005
	amoA-1F	GGGGTTTCTACTGGTGGT	56 °C	Betaproteobacterial <i>amoA</i>	Rotthauwe et al., 1997
	AmoA-RNEW	CCCCTCBGSAAAVCCTCTTC	56 °C	Betaproteobacterial <i>amoA</i>	Rotthauwe et al., 1997
qPCR	Lake 386F	ACGACWRGGTYARCCGAGTGRTT	60 °C	Thaumarchaeal 16S rRNA	This study
	957R	CGGCGTTGACTCCAATTG	60 °C	Thaumarchaeal 16S rRNA	Ochsenreiter et al., 2003
	Arch-amoA-for	CTGAYTGGGCYTGGACATC	58.5 °C	Thaumarchaeal <i>amoA</i>	Wuchter et al., 2006
	Arch-amoA-rev	TTCTTCTTTGTTGCCAGTA	58.5 °C	Thaumarchaeal <i>amoA</i>	Wuchter et al., 2006
	amoA-1F	GGGGTTTCTACTGGTGGT	56 °C	Betaproteobacterial <i>amoA</i>	Rotthauwe et al., 1997
	AmoA-RNEW	CCCCTCBGSAAAVCCTCTTC	56 °C	Betaproteobacterial <i>amoA</i>	Rotthauwe et al., 1997

of an initial denaturing step for 5 min at 95 °C; all the programs were followed by 45 cycles consisting of 94 °C for 30 s, specific annealing temperature for 60 s and 72 °C for 60 s. The fluorescence signal was read in each cycle after the elongation step to ensure stringent product quantification. Specificity for the qPCR reaction was tested by melting curve analysis (60 °C–94 °C) in order to identify non-specific PCR products such as primer dimers or fragments with unexpected lengths and then by electrophoresis on 1% agarose gel (data not shown). Standard curves for the target genes (thaumarchaeal 16S rRNA, AOA or AOB *amoA* genes) were generated from a mix of clones representative of the environments studied. Next, serial dilutions of previously titrated suspensions were performed and amplified by conventional PCR from environmental clones and then purified (QIAquick, Qiagen) and quantified in this study. All reactions were performed with standard curves spanning from 10^1 to 10^8 copies per μ L. Mean PCR efficiencies and correlation coefficients for standard curves were as follows: for the thaumarchaeal 16S rRNA assay, 111%, $r^2 = 0.96$, for the archaeal *amoA* assay, 111%, $r^2 = 0.98$ and for the betaproteobacterial *amoA* assay, 69.8%, $r^2 = 1.00$.

2.4. 16S rRNA and *amoA* gene libraries

For each lake, 16S rRNA and AOA *amoA* genes were PCR-amplified from a sample corresponding to homothermy and thermal stratification periods using archaeal-specific primers (Table 2) and conditions previously described (Francis et al., 2005; Jurgens et al., 2000; Rotthauwe et al., 1997). PCR products were cloned using a TOPO TA cloning kit according to the manufacturer's instructions (Invitrogen). Cloned inserts were PCR-amplified using the M13 forward and reverse primers, and amplicons were digested with restriction endonucleases HaeIII and RsaI (Qbiogene) to screen 16S rRNA and *amoA* libraries, respectively. To ensure correspondence between the OTU and RFLP pattern, at least two clones of each OTU were selected for sequencing. Sequencing reactions were performed by MWG (<http://www.mwg-biotech.com>). The archaeal 16S rRNA gene sequences from Lake Bourget obtained in this study have been archived in GenBank under accession numbers EU490289–EU490317; for Lake Aydat, JF980343–JF980358

and JF980377–JF980384; and for the Sep reservoir: JF980359–JF980376. The archaeal *amoA* sequences determined in this study have been deposited in the GenBank database under accession numbers JN089884–JN089917.

2.5. Phylogenetic and statistical analyses

16S rRNA sequences were screened for potential chimeric structures using the Bellerophon program (Huber et al., 2004). Sequences aligned with “Greengene” (<http://greengenes.lbl.gov>) were used for phylogenetic analyses on ARB software (Ludwig et al., 2004) using a Greengene database enriched with lacustrine archaeal sequences. Partial sequences were inserted into the ARB tree using the parsimony method. The resulting tree was pruned to retain the closest relative's sequences and highlighted the main environmental archaeal clades in aquatic ecosystems. The *amoA* sequences from the three lakes were blasted against the NR protein database (tblastx – e-value $\leq 10^{-4}$). Corresponding nucleotide sequences were retrieved and further refined for sequences exhibiting at least 97% identity over a minimum 90% of query sequence length. This *amoA* database was further enriched with reference sequences found in lakes and not retrieved by BLAST (Francis et al., 2005; Nicol and Schleper, 2006; Prosser and Embley, 2002). Sequence alignment was performed using MUSCLE (Edgar, 2004). The maximum likelihood tree was constructed with PhyML (Guindon and Gascuel, 2003), using the GTR + I + G evolutionary model and 100 bootstrap replicates. Good's coverage, richness estimators and statistical tests were performed using the R package and Vegan (<http://cran.r-project.org/web/packages/vegan/index.html>).

Relations between the relative abundances of *Thaumarchaeota*, archaeal and bacterial *amoA* genes and environmental factors (temperature, Chl_a, nitrate, ammonium, phosphate, and oxygen) were examined by redundancy analysis (RDA) performed in XL Stat 2010 (Addinsoft, Paris). This RDA considered changes in thaumarchaeal and ammonia-oxidizing community genes abundances on Lake Bourget and the Sep reservoir in relation to environmental parameters.

Unifrac (<http://bmf2.colorado.edu/unifrac/index.psp>) metrics was used for comparing microbial communities based on phylogenetic information (Lozupone and Knight, 2005).

3. Results

3.1. Changes in ammonia-oxidizing microorganisms between homothermy and thermal stratification periods

The main physical, chemical and biological characteristics of the three studied lakes are listed in Table 1. The three lakes were very different, especially due to their catchment area, maximum depth and nutrient content. To better understand the dynamics and potential contribution of AOA and AOB to

ammonia oxidation, their abundances were examined by quantitative PCR (Fig. 1) during both homothermy (H) and thermal stratification (TS). In the epilimnion of Lake Bourget, AOA and AOB communities were more abundant during homothermy with, respectively, 64 and 3.6 times more *amoA* gene copies than during thermal stratification (Fig. 1A). Moreover, this ecosystem seemed to be characterized by a shift in ammonia-oxidizing communities illustrated by the predominance of AOA over AOB during winter, in contrast to thermal stratification dominated by AOB. In the Sep reservoir

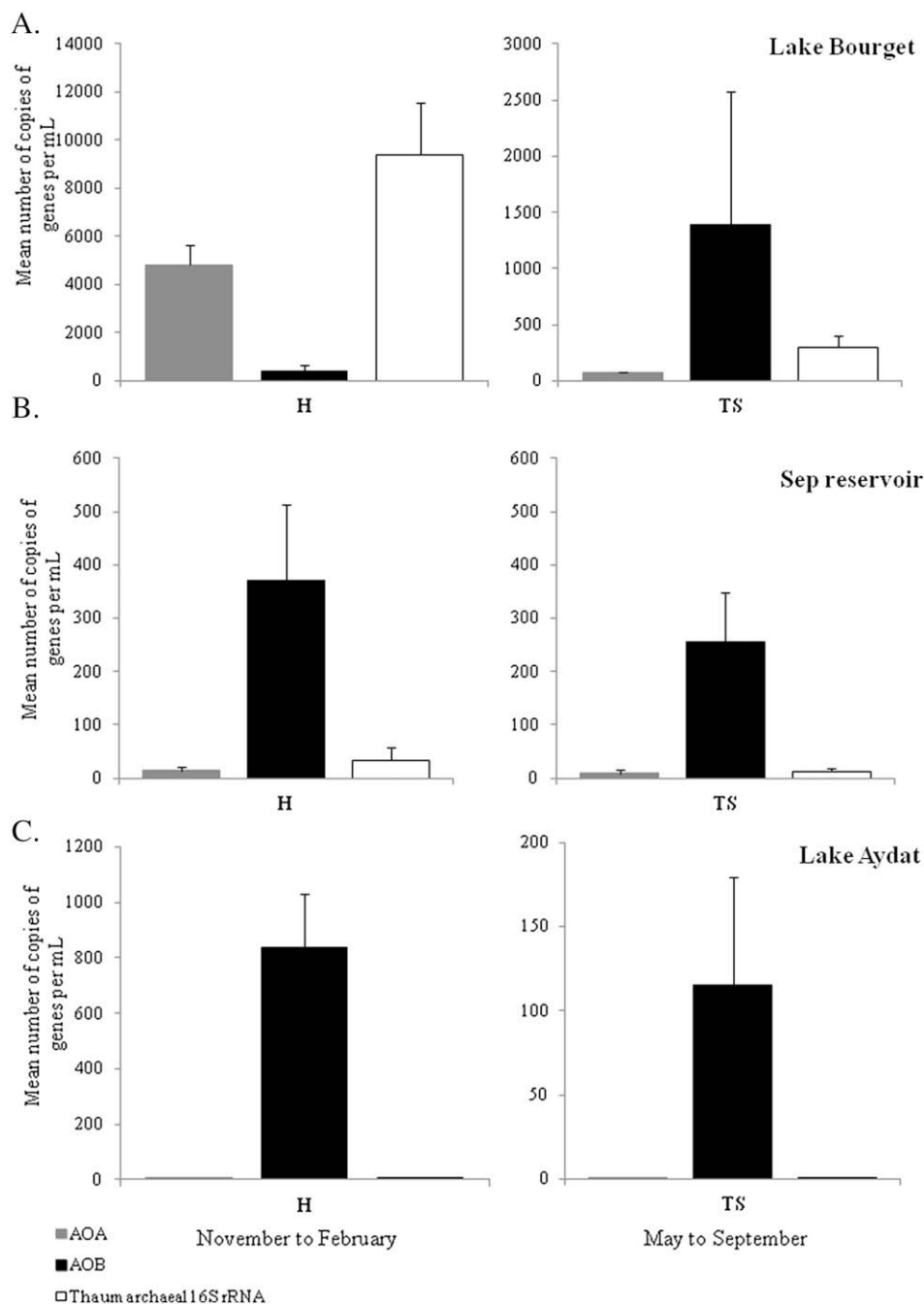


Fig. 1. Abundance of archaeal 16S rRNA, AOA and AOB *amoA* genes in the epilimnion of Lake Bourget (A), the Sep reservoir (B) and Lake Aydat (C). Mean values were calculated on sampling dates during winter homothermy, from November to February (H) and summer thermal stratification, from May to September (TS).

epilimnion, thaumarchaeal 16S rRNA and *amoA* genes were retrieved during both periods (Fig. 1B) and were, respectively, 1.5 and 2.7 times more abundant during homothermy than during thermal stratification. Interestingly, in this ecosystem AOB were always predominant, about 26 and 28 times more abundant than AOA during homothermy and thermal stratification, respectively. No significant number of archaeal *amoA* or thaumarchaeal 16S rRNA genes could be detected in Lake Aydat, while AOB genes were consistently present and most abundant during homothermy (Fig. 1C). In Lake Bourget and the Sep reservoir, where both ammonia-oxidizing communities were retrieved, RDA was performed to evaluate the potential link between those populations and environmental parameters. The first two axes explained 99.77% of variance in species abundance (AOA, AOB and *Thaumarchaeota*, Fig. 2). This assessed the positive link between the abundance of AOA, AOB and *Thaumarchaeota* and the oxygen concentration in the epilimnion of these ecosystems. All Sep reservoir sampling dates (H and TS) were related to nutrients and Chl *a* content, while in Lake Bourget, only thermal stratification points were related to these parameters (Fig. 2).

3.2. Phylogenetic analysis of archaeal 16S rRNA genes

To track putative changes in archaeal community composition, phylogenetic analyses were conducted during both periods chosen: H and TS. Based on their phylogenetic affiliation, archaeal 16S rRNA sequences were distributed among *Euryarchaeota*, *Thaumarchaeota* and *Crenarchaeota*, with 35, 32 and 1 OTUs, respectively, in the 3 lakes (Table 3 and see Supplementary Fig. 1A for phylogenetic analysis). All thaumarchaeal sequences retrieved from the three lakes fell into clades I.1a, SAGMCG-1 and I.1b. Nine OTUs isolated exclusively from the Sep reservoir clustered together within the SAGMCG-1 group. Twenty-two thaumarchaeal OTUs originating from the studied lakes branched within

clade I.1a. Interestingly, these sequences clustered within two subclades mainly composed of sequences isolated from lake water, rhizosphere and sediment. Subcluster A included 5 sequences from the Sep reservoir (H), Lake Bourget (TS) and Aydat (H). Subcluster B comprised sequences from Lake Aydat and Bourget. UniFrac analysis based on thaumarchaeal 16S phylogeny revealed that the majority of the tested environments were significantly different (Table 4A, p -value < 0.05), like Sep H and Sep TS or Aydat H and Aydat TS. The greatest distances (>0.70) retrieved in that dataset involved the Sep reservoir, indicating that thaumarchaeal communities associated with this ecosystem were different (Table 4A). This was supported by the size of the SAGMCG-1 group found in this lake, not even retrieved in the other lakes (Table 3). *Euryarchaeota* sequences affiliated with Methanosaetaceae (14 OTUs), Methanomicrobiaceae (15 OTUs) and WHD3-02 (3 OTUs) were detected in all lakes, while other environmental clades were restricted to a single lake, e.g. cluster II in Lake Bourget and PMC1 in the Sep reservoir (Table 3).

3.3. Phylogenetic analysis of archaeal *amoA* genes

As all *amoA* gene sequences from AOB were close to *Nitrosomonas europaea* (data not shown), the analysis focused on archaeal *amoA* diversity, which was analyzed from the same water samples as for 16S rRNA diversity analyses, during homothermy and thermal stratification. *amoA* genes were detected only during homothermy, and richness, expressed by the S_{chao1} index, ranged from 8.75 (Sep) to 21.5 (Bourget) (Table 3). UniFrac analysis based on *amoA* phylogenies did not show significant differences between the three studied lakes (Table 4B). The largest distance retrieved was between Lake Aydat and Lake Bourget (0.72, $p = 0.07$), which differed most in terms of nutrient content (Table 1). All OTUs were grouped into the two major clusters classified as “marine/sediment” cluster (cluster A) and “soil/sediment” cluster (cluster B) by Francis et al. (2005) (Fig. 3A). Most of them fell into 3 major subclusters of cluster A (A1–A3) and were more closely related to *amoA* sequences from sediment than from water columns. For the three lakes, half of the OTUs were in the subcluster “freshwater and low salinity A1”, with A1.1, a well-supported clade, comprising sequences retrieved from all studied lakes and A1.2 containing only three OTUs from Lake Bourget (Fig. 3B). In addition, the subcluster “freshwater lake and hot spring A3” comprised up to 37.5% of all OTUs. Only two OTUs from Lake Aydat and one OTU from the Sep reservoir fell into “soil/sediment” cluster B (Fig. 3A). Within this cluster, subcluster B1, including Aydat and Sep sequences, was particularly rich in sequences from miscellaneous environments as waste waters, estuaries or soils (Inagaki et al., 2003).

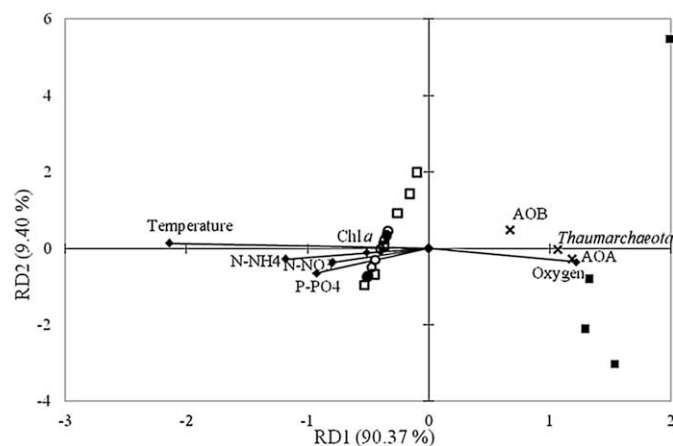


Fig. 2. RDA showing correlation between thaumarchaeal 16S rRNA and ammonia-oxidizing community abundances and selected environmental variables (temperature, oxygen, NH_4 , NO_3 and PO_4 concentrations). Homothermy (■) and thermal stratification (□) for Lake Bourget and Sep reservoir homothermy (●) and thermal stratification (○).

4. Discussion

Planktonic freshwater habitats have emerged as one of the largest reservoirs of archaeal diversity (Auguet et al., 2011;

Table 3

Diversity of clone libraries for 16S rRNA and *amoA* gene in the epilimnion of Lake Bourget, the Sep reservoir and Aydat.

Lakes	Clade	Bourget		Sep		Aydat	
Sampling period		TS	H	TS	H	TS	H
16S rRNA gene							
no° of clones		57	60	60	35	39	37
n° of OTUs		21	5	9	9	8	16
S _{chao1}		87	5	10	9.3	13	18.5
Cgood's		0.79	0.98	0.95	0.97	0.87	0.81
<i>Thaumarchaeota</i>	I.1a	8 (33)	5 (60)		1 (1)	7 (38)	1 (2)
	I.1b						1 (1)
	SAGMCG-1			5 (37)	4 (18)		
<i>Crenarchaeota</i>	Non-affiliated				1 (52)		
<i>Euryarchaeota</i>	Methanomicrobiaceae	4 (4)		2 (17)	1 (2)	1 (1)	7 (15)
	Methanosaetaceae	6 (15)		1 (2)	1 (6)		6 (17)
	WHD3-02	1 (1)		1 (1)			1 (1)
	DHVE-5	1 (1)					
	Cluster II	1 (3)					
	PMC1				1 (1)		
<i>amoA</i> gene							
n° of clones			31		23		49
n° of OTUs			11		8		15
S _{chao1}			21.5		8.75		16.5
Cgood's			0.97		0.9		0.91
	Freshwater and low salinity lake A1.1		5 (24)		2 (3)		5 (21)
	Ubiquitous A1.2		3 (3)				
	Hot spring and lake A2						2 (2)
	Freshwater lake and hot spring A3		3 (4)		5 (19)		6 (19)
	Soil/sediment B				1 (1)		2 (7)

Main archaeal 16S rRNA and *amoA* clades in the three studied lakes, with number of OTUs (number of clones). H: homothermy period; TS: thermal stratification period. S_{chao1}: non-parametric richness estimator and Cgood's coverage value.

Lliros et al., 2010); however, the contribution of ammonia-oxidizing *Archaea* to the nitrogen cycle of these ecosystems is not well characterized. This study showed that AOB were predominant in nutrient-rich conditions (Sep reservoir

and Lake Aydat), while AOA dominated where ammonia values were the lowest (homothermy period in Lake Bourget). Seasonal changes in the abundance of AOA have been observed in surface waters of oligotrophic mountain lakes, with an increase in summer in correlation with the dynamics of ammonium and nitrite concentrations (August et al., 2011), but their relative contribution compared to AOB remained unknown. In this study, we also found seasonal changes in AOA abundance in Lake Bourget; however, contrary to August et al. (2011), the increase in AOA abundance was observed during winter, as previously observed in several marine ecosystems (Galand et al., 2010; Wuchter et al., 2006). Indeed, Galand et al. (2010) showed that *Thaumarchaeota* group I abundance was highest during winter and significantly correlated with archaeal ammonia monooxygenase (*amoA*) gene copies and nitrite concentrations in the Mediterranean Sea. It has been suggested that winter conditions are favorable to AOA, facilitating competition with phytoplankton (Pitcher et al., 2011), while stratification in the epilimnetic zone could create unfavorable conditions for *Archaea* (competition, nutrient availability, (Winter et al., 2009). In this study, we investigated the relative abundance of AOA and AOB in relation to nutrient concentrations and for two different environmental periods of the year: summer thermal stratification and winter homothermy. Thermal stratification leads to an upper oxic layer separated from a deeper anoxic layer by a marked chemocline, contrasting

Table 4

Unifrac distances for comparing thaumarchaeal communities (A) according to the lake and the period (TS or H).

A.						
	Aydat — TS	Bourget — TS	Sep — TS	Aydat — H	Bourget — H	Sep — H
Aydat — TS	0	0.27	0.00	0.02	0.44	0.00
Bourget — TS	0.23	0	0.00	0.06	0.56	0.00
Sep — TS	0.78	0.80	0	0.04	0.01	0.05
Aydat — H	0.48	0.47	0.82	0	0.06	0.69
Bourget — H	0.14	0.15	0.78	0.47	0	0.00
Sep — H	0.48	0.45	0.47	0.53	0.47	0
B.						
	Aydat		Bourget		Sep	
Aydat	0		0.07		0.37	
Bourget	0.72		0		0.33	
Sep	0.45		0.65		0	

Values shown in the left panel (grey) correspond to Unifrac distances and, in the right panel, to probabilities. (B) Unifrac distances for comparing *amoA* phylogeny retrieved during winter homothermy according to the lake. Values in the left panel (grey) correspond to Unifrac distances and, in the right panel, to probabilities.

with the uniform conditions found throughout the water column during winter homothermy and providing potential different niches associated with different communities. For instance, several studies have shown maximal abundance of AOA in the oxycline of stratified lakes (Auguet et al., 2011; Lliros et al., 2010; Pouliot et al., 2009). However, in our study, no enrichment of the archaeal community (16S rRNA or *amoA*) was observed in the oxycline (data not shown) and we therefore focused on the surface layer. The dominance of AOB when nutrient content was highest strengthens results obtained in specific soils showing significantly higher values of AOB than AOA (Boyle-Yarwood et al., 2008). In contrast, it has also been observed that oligotrophic waters constitute an ideal habitat for AOA, since in situ ammonia concentrations were at the lower end of AOB affinity ($>1 \mu\text{M}$) (Bollmann et al., 2002) and AOA may easily outcompete AOB and heterotrophs under ammonia limitations (Martens-Habben et al., 2009). Furthermore, in soils, the relative abundance of AOA and AOB generally changed with depth, with the AOA number remaining relatively constant while AOB number dramatically decreased (Di et al., 2010; Leininger et al., 2006), indicating that AOA may be particularly well adapted to conditions of low levels of available nutrients. Similarly, AOA appeared to be both numerically and functionally more active compared to AOB in marine environments (Mincer et al., 2007; Ward et al., 2007; Wuchter et al., 2006). Nutrient availability could therefore be an important factor in controlling AOA/AOB ratios in lakes, as suggested in our study, with a predominance of AOA over AOB in Lake Bourget during winter when ammonia values are lowest.

AOA dynamics and their implication in the nutrient cycle of lakes might also be related to the structure of *Thaumarchaeota* inhabiting these different lakes. To investigate the potential rearrangement of archaeal populations in the different lakes throughout the year, we analyzed both archaeal 16S rRNA and *amoA* diversity. In our study, we always found more *Thaumarchaeota* (i.e. 16S rRNA abundance) than AOA. Since the complete genomes of mesophilic *Archaea* showed one copy of 16S rRNA and *amoA* (Könneke et al., 2005), those results suggested that *Thaumarchaeota* were not consistently involved in the nitrogen cycle and could be represented by different populations. However, we cannot exclude that detection of archaeal *amoA* may be biased by both detection limits and the different sets of primers used in this study. Phylotypes retrieved in this study are divided into two groups, one consisting of cosmopolitan phylotypes and one composed of more regionally restricted populations. Among the thaumarchaeal cosmopolitan phylotypes, clade I.1a (or MGI) was retrieved and well represented in the three lakes studied here. This clade, I.1a, has also been well described in meromictic Lake Kivu (Lliros et al., 2010) and several alpine lakes (Auguet et al., 2011). Group I.1b was also present, as previously shown, in Lake Kivu (Lliros et al., 2010). This group was typically associated with sequences originating from soils (Bintrim et al., 1997), but tended to be retrieved from aquatic ecosystems (Lliros et al., 2010). On

the other hand, some phylotypes were restricted to one lake, as shown by the predominance of the thaumarchaeal SAGMCG-1 clade in the Sep reservoir during homothermy and thermal stratification. The SAGMCG-1 group has previously been encountered in diverse environments such as hot springs, and plant roots (Nicol and Schleper, 2006) and its sole presence in the Sep reservoir could result from the reservoir's extensive water renewal (Boucher et al., 2006), probably leading to terrestrial input. Indeed, this reservoir has been built to provide agricultural water and its regular draining could provide particular environmental parameters (nutrient availability) affecting community growth. The presence of several thaumarchaeal OTUs affiliated with the SAGMCG-1 clade in the Sep reservoir, associated with the *amoA* richness in this ecosystem, also leads us to speculate that this cluster could represent an important player in ammonia oxidation in some ecosystems, as previously shown in draining mines (Pester et al., 2011). This particular distribution could provide insights into new sampling strategies for elucidating phylogenetic history but also archaeal metabolic functions. The distribution of phylotypes in cosmopolitan and more restricted groups raises questions about biotic and abiotic processes and the balance between the two, both of which are responsible for structuring these populations and determining their patterns of distribution. The richness of lacustrine archaeal *amoA* genes found in our study exceeded that of *amoA* genes at the different sites sampled thus far, including marine waters and freshwater ecosystems (Agogué et al., 2008; Auguet et al., 2009; Francis et al., 2005; Pouliot et al., 2009). It also revealed differences between the studied lakes, but also substantial congruence with previous phylogenies (Herrmann et al., 2009; Mosier and Francis, 2008; Prosser and Nicol, 2008) and assessed the presence of the “freshwater and low salinity” clade (A1) also described in Hu et al. (2010) and Auguet et al. (2011). Around 40% of *amoA* OTUs extracted from the three lakes fell into cluster A3 comprised almost entirely (85%) of hot springs, rhizospheres and sediment of freshwater macrophytes and freshwater lake sequences. This clade highlighted an unexpected similarity between freshwater and hot spring ammonia-oxidizing thaumarchaeal communities. These results were particularly interesting, as no known AOB had ever been found in hot springs (Takai et al., 2001).

Overall, *Archaea* are common components of freshwater plankton and tend to occupy different ecological niches, although their metabolism and physiological roles remain challenging. In this study, we show that nutrient availability is one of the main environmental factors involved in both AOA and AOB community establishment. Our results also suggest potential rearrangements of *Thaumarchaeota* in diverse populations exhibiting differing physiological traits and varying with time and space, suggesting the presence of different ecotypes. However, this preliminary study needs to be completed by a more extensive spatio-temporal survey, AOA transcript abundance and bulk nitrification measurements, but also by precise characterization of the putative thaumarchaeal ecotypes through single cell approaches.



Fig. 3. Phylogenetic tree of archaeal *amoA* sequences recovered from Lake Bourget, the Sep reservoir and Lake Aydat (A). Environmental sequences are shown in bold. Bootstrap values (>75%) are indicated at each branch point. Scale bar: 0.1 estimated number of substitutions per nucleotide position. (B) Detailed view of the phylogenetic tree of archaeal *amoA* sequences retrieved in the "freshwater and low salinity A1" clade. Environmental sequences are shown in bold. Bootstrap values (>75%) are indicated at each branch point. Scale bar: 0.05 estimated number of substitutions per nucleotide position.



Fig. 3. (Continued)

Acknowledgments

This work is being supported by CNRS Program EC2CO. Physicochemical data was performed and analyzed by the observatory on peri-alpine lakes (INRA, UMR CARTEL, F-74200 Thonon Les Bains). We thank G. Paolini and P. Perney for technical contributions to sampling and analysis of physical and chemical parameters in Lake Bourget and the ATT Society for helping to improve the English text.

Appendix A. Supplementary material

Supplementary material related to this article can be found at <http://dx.doi.org/10.1016/j.resmic.2013.01.004>.

References

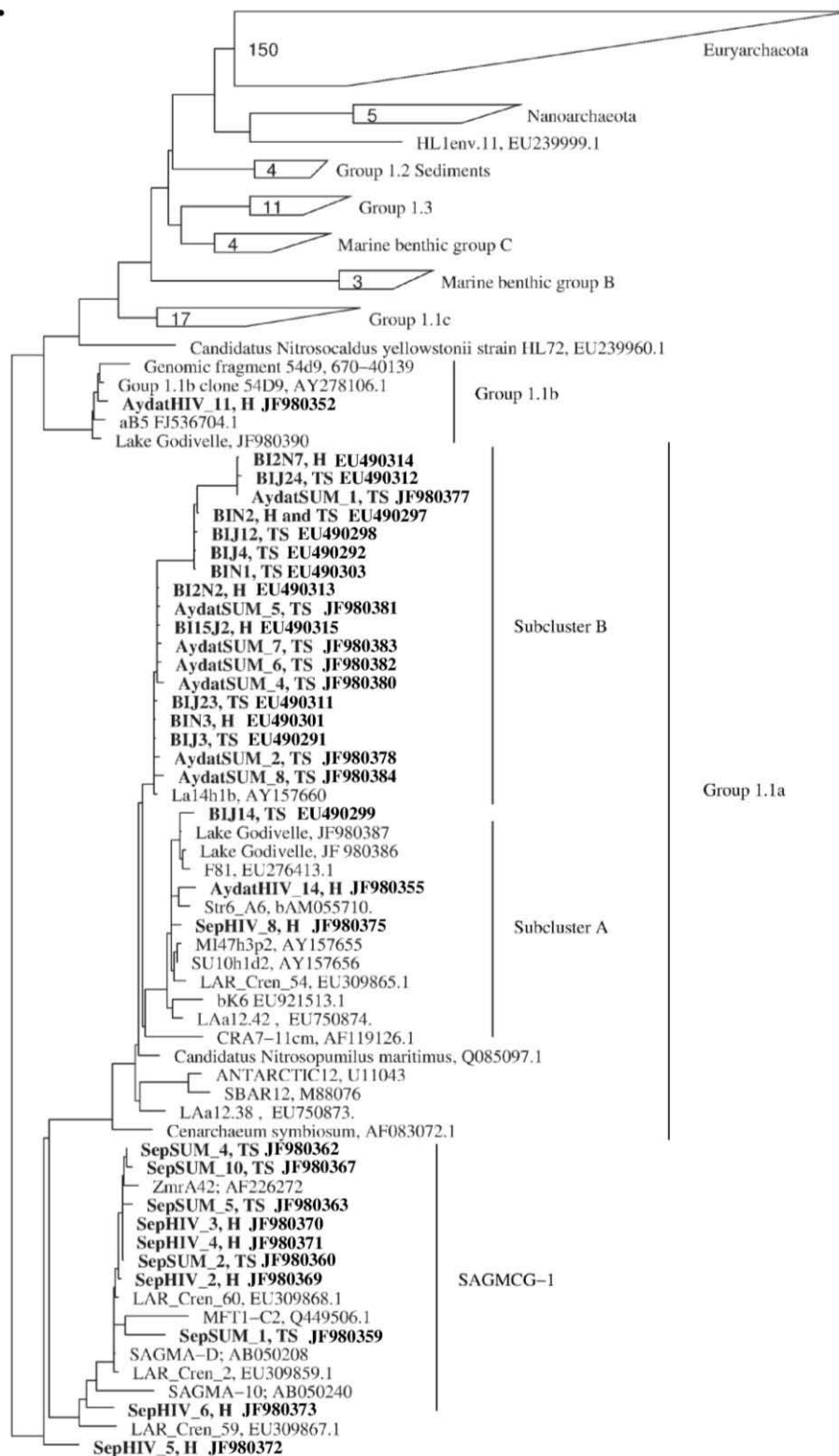
- Agogué, H., Brink, M., Dinasquet, J., Herndl, G.J., 2008. Major gradients in putatively nitrifying and non-nitrifying *Archaea* in the deep North Atlantic. *Nature* 456, 788–791.
- Auguet, J.C., Barberan, A., Casamayor, E.O., 2009. Global ecological patterns in uncultured *Archaea*. *ISME J.* 4, 182–190.
- Auguet, J.C., Casamayor, E.O., 2008. A hotspot for cold *Crenarchaeota* in the neuston of high mountain lakes. *Environ. Microbiol.* 10, 1080–1086.
- Auguet, J.C., Nomokonova, N., Camarero, L., Casamayor, E.O., 2011. Seasonal changes of freshwater ammonia-oxidizing archaeal assemblages and nitrogen species in oligotrophic alpine lakes. *Appl. Environ. Microbiol.* 77, 1937–1945.
- Bintrim, S.B., Donohue, T.J., Handelsman, J., Roberts, G.P., Goodman, R.M., 1997. Molecular phylogeny of *Archaea* from soil. *Proc. Natl. Acad. Sci. U.S.A.* 94, 277–282.
- Bollmann, A., Bar-Gilissen, M.J., Laanbroek, H.J., 2002. Growth at low ammonium concentrations and starvation response as potential factors involved in niche differentiation among ammonia-oxidizing *bacteria*. *Appl. Environ. Microbiol.* 68, 4751–4757.
- Boucher, D., Richardot, M., Thenot, A., Debroas, D., 2006. Incorporation of 3H-thymidine by different prokaryotic groups in relation to temperature and nutrients in a lacustrine ecosystem. *Microb. Ecol.* 52, 399–407.
- Boyle-Yarwood, S.A., Bottomley, P.J., Myrold, D.D., 2008. Community composition of ammonia-oxidizing *bacteria* and *archaea* in soils under stands of red alder and Douglas fir in Oregon. *Environ. Microbiol.* 10, 2956–2965.
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., Forterre, P., 2008. Mesophilic *crenarchaeota*: proposal for a third archaeal phylum, the *Thaumarchaeota*. *Nat. Rev. Microbiol.* 6, 245–252.
- DeLong, E.F., 1992. *Archaea* in coastal marine environments. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5685–5689.
- Di, H.J., Cameron, K.C., Shen, J.P., Winefield, C.S., O'Callaghan, M., Bowatte, S., He, J.Z., 2010. Ammonia-oxidizing *bacteria* and *archaea* grow under contrasting soil nitrogen conditions. *FEMS Microbiol. Ecol.* 72, 386–394.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.
- Erguder, T.H., Boon, N., Wittebolle, L., Marzorati, M., Verstraete, W., 2009. Environmental factors shaping the ecological niches of ammonia-oxidizing *archaea*. *FEMS Microbiol. Rev.* 33, 855–869.
- Francis, C.A., Beman, J.M., Kuypers, M.M., 2007. New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J.* 1, 19–27.
- Francis, C.A., Roberts, K.J., Beman, J.M., Santoro, A.E., Oakley, B.B., 2005. Ubiquity and diversity of ammonia-oxidizing *Archaea* in water columns and sediments of the ocean. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14683–14688.
- Galand, P.E., Gutierrez-Provecho, C., Massana, R., Gasol, J., Casamayor, E.O., 2010. Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea. *Limnol. Oceanogr.* 55, 2117–2125.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Hallam, S.J., Konstantinidis, K.T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., Preston, C., de la Torre, J., Richardson, P.M., DeLong, E.F., 2006. Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18296–18301.
- Hatzenpichler, R., Lebedeva, A., Spieck, E., Stoecker, K., Richter, A., Daims, H., Wagner, M., 2008. A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2134–2139.
- Herrmann, M., Saunders, A.M., Schramm, A., 2009. Effect of lake trophic status and rooted macrophytes on community composition and abundance of ammonia-oxidizing prokaryotes in freshwater sediments. *Appl. Environ. Microbiol.* 75, 3127–3136.
- Hu, A., Yao, T., Jiao, N., Liu, Y., Yang, Z., Liu, X., 2010. Community structures of ammonia-oxidizing *archaea* and *bacteria* in high-altitude lakes on the Tibetan Plateau. *Freshwater Biol.* 55, 2375–2390.
- Huber, T., Faulkner, G., Hugenholtz, P., 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20, 2317–2319.
- Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K., Nealson, K.H., Horikoshi, K., 2003. Microbial communities associated with geological horizons in coastal seafloor sediments from the sea of Okhotsk. *Appl. Environ. Microbiol.* 69, 7224–7235.
- Jardillier, L., Boucher, D., Personnic, S., Jacquet, S., Thenot, A., Sargos, D., Amblard, C., Debroas, D., 2005. Relative importance of nutrients and mortality factors on prokaryotic community composition in two lakes of different trophic status: microcosm experiments. *FEMS Microbiol. Ecol.* 53, 429–443.
- Jurgens, G., Glockner, F., Amann, R., Saano, A., Montonen, L., Likolammi, M., Munster, U., 2000. Identification of novel *Archaea* in bacterioplankton of a boreal forest lake by phylogenetic analysis and fluorescent in situ hybridization. *FEMS Microbiol. Ecol.* 34, 45–56.
- Karner, M.B., DeLong, E.F., Karl, D.M., 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409, 507–510.
- Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., Stahl, D.A., 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437, 543–546.
- Kowalchuk, G.A., Stephen, J.R., 2001. Ammonia-oxidizing bacteria: a model for molecular microbial ecology. *Annu. Rev. Microbiol.* 55, 485–529.
- Lefranc, M., Thenot, A., Lepere, C., Debroas, D., 2005. Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl. Environ. Microbiol.* 71, 5935–5942.
- Leininger, S., Urich, T., Schlöter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J.I., Schuster, S.C., Schleper, C., 2006. *Archaea* predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442, 806–819.
- Lliros, M., Gich, F., Plasencia, A., Auguet, J.C., Darchambeau, F., Casamayor, E.O., Descy, J.P., Borrego, C., 2010. Vertical distribution of ammonia-oxidizing *crenarchaeota* and methanogens in the epipelagic waters of Lake Kivu (Rwanda-Democratic Republic of the Congo). *Appl. Environ. Microbiol.* 76, 6853–6863.
- Lorenzen, C.J., 1967. Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnol. Oceanogr.* 12, 343–346.
- Lozupone, C., Knight, R., 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71, 8228–8235.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., et al., 2004. ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371.
- Martens-Habbena, W., Berube, P.M., Urakawa, H., de la Torre, J.R., Stahl, D.A., 2009. Ammonia oxidation kinetics determine niche separation of nitrifying *Archaea* and *Bacteria*. *Nature* 461, 976–979.
- Mincer, T.J., Church, M.J., Taylor, L.T., Preston, C., Karl, D.M., DeLong, E.F., 2007. Quantitative distribution of presumptive archaeal and bacterial

- nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ. Microbiol.* 9, 1162–1175.
- Mosier, A.C., Francis, C.A., 2008. Relative abundance and diversity of ammonia-oxidizing *archaea* and bacteria in the San Francisco Bay estuary. *Environ. Microbiol.* 10, 3002–3016.
- Nicol, G.W., Schleper, C., 2006. Ammonia-oxidising *Crenarchaeota*: important players in the nitrogen cycle? *Trends Microbiol.* 14, 207–212.
- Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, L., Schleper, C., 2003. Diversity and abundance of *Crenarchaeota* in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environ. Microbiol.* 5, 787–797.
- Park, H.D., Wells, G.F., Bae, H., Criddle, C.S., Francis, C.A., 2006. Occurrence of ammonia-oxidizing archaea in wastewater treatment plant bioreactors. *Appl. Environ. Microbiol.* 72, 5643–5647.
- Pelletier, J.P., Orand, A., 1978. Appareil de prélèvement d'un échantillon dans un fluide. Brevet d'intention 76, 08579.
- Pester, M., Schleper, C., Wagner, M., 2011. The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* 14, 300–306.
- Pitcher, A., Villanueva, L., Hopmans, E.C., Schouten, S., Reichart, G.J., Sinninghe Damste, J.S., 2011. Niche segregation of ammonia-oxidizing *archaea* and anammox *bacteria* in the Arabian Sea oxygen minimum zone. *ISME J.* 5, 1896–1904.
- Pouliot, J., Galand, P.E., Lovejoy, C., Vincent, W.F., 2009. Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two meromictic High Arctic lakes. *Environ. Microbiol.* 11, 687–699.
- Prosser, J.I., Embley, T.M., 2002. Cultivation-based and molecular approaches to characterisation of terrestrial and aquatic nitrifiers. *Antonie Van Leeuwenhoek* 81, 165–179.
- Prosser, J.I., Nicol, G.W., 2008. Relative contributions of *archaea* and bacteria to aerobic ammonia oxidation in the environment. *Environ. Microbiol.* 10, 2931–2941.
- Rotthauwe, J.H., Witzel, K.P., Liesack, W., 1997. The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl. Environ. Microbiol.* 63, 4704–4712.
- Schleper, C., Jurgens, G., Jonscheit, M., 2005. Genomic studies of uncultivated *archaea*. *Nat. Rev. Microbiol.* 3, 479–488.
- Strickland, J., Parsons, T., 1968. A Practical Handbook of Sea Water Analysis.
- Takai, K., Komatsu, T., Inagaki, F., Horikoshi, K., 2001. Distribution of *archaea* in a black smoker chimney structure. *Appl. Environ. Microbiol.* 67, 3618–3629.
- Treusch, A.H., Leininger, S., Kletzin, A., Schuster, S.C., Klenk, H.P., Schleper, C., 2005. Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic *crenarchaeota* in nitrogen cycling. *Environ. Microbiol.* 7, 1985–1995.
- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., et al., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66–74.
- Ward, B., Capone, D.G., Zehr, J.P., 2007. What's new in the nitrogen cycle? *Oceanography* 20, 101–109.
- Winter, C., Kerros, M.E., Weinbauer, M.G., 2009. Seasonal changes of bacterial and archaeal communities in the dark ocean: evidence from the Mediterranean Sea. *Limnol. Oceanogr.* 54, 160–170.
- Wuchter, C., Abbas, B., Coolen, M.J., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., Schouten, S., Sinninghe Damste, J.S., 2006. Archaeal nitrification in the ocean. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12317–12322.

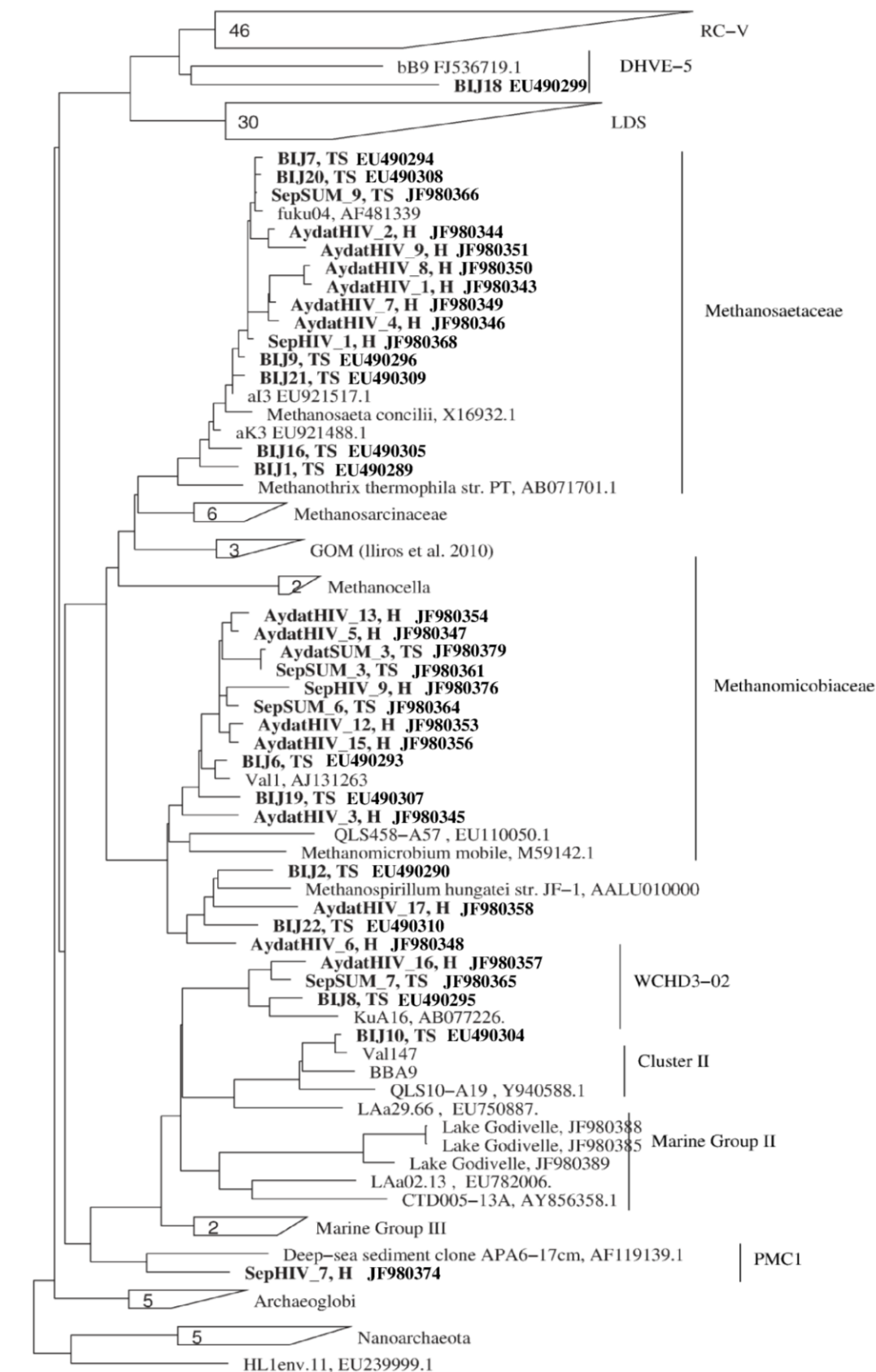
Supplementary material for online submission

Supplementary Fig. 1

A.



0.10

B.

0.10

5.3 Article 4 : Temporal and spatial dynamics of active prokaryotes nitrifiers and archaeal communities from river to sea

Mylène Hugoni, Hélène Agogué, Najwa Taib, Isabelle Domaizon, Anne Moné, Pierre E. Galand, Gisèle Bronner, Didier Debroas, Isabelle Mary
En préparation

Temporal and spatial dynamics of active prokaryotes nitrifiers and archaeal communities from river to sea

Mylène Hugoni^{1,2}, Hélène Agogué³, Najwa Taib^{1,2}, Isabelle Domaizon⁴, Anne Moné^{1,2}, Pierre
5 E. Galand^{5,6}, Gisèle Bronner^{1,2}, Didier Debroas^{1,2}, Isabelle Mary^{1,2}

(1) Clermont Université, Université Blaise Pascal, Laboratoire "Microorganismes : Génome et
Environnement", BP 10448, F-63000 CLERMONT-FERRAND, France;

(2) CNRS, UMR 6023, LMGE, F-63171 AUBIERE, France;

10 (3) Littoral, Environnement et Sociétés, UMR 7266, CNRS, University of La Rochelle,
17000 La Rochelle, France;

(4) Institut National de la Recherche Agronomique, UMR 42 Centre Alpin de Recherche sur
les Réseaux Trophiques et Ecosystèmes Limniques, F-74200 Thonon les Bains, France;

(5) Université Pierre et Marie Curie–Paris 6, UMR 8222, Laboratoire d’Ecogéochimie des
15 Environments Benthiques (LECOB), UMR 7621, Laboratoire d’Océanographie Microbienne
(LOMIC), Observatoire Océanologique, F-66650 Banyuls-sur-Mer, France;

(6) CNRS, UMR 8222, LECOBI, Observatoire Océanologique, F-66650 Banyuls-sur-Mer,
France.

20

Running title: Nitrifiers’ dynamics in an estuarine gradient

Subject Category: Microbial population and community ecology

Keywords: Ammonia oxidation / *amoA* / *Archaea* / gradient / diversity

25

Abstract

Nitrification plays an important role in estuarine biogeochemistry, however, the activity and the diversity of Ammonia Oxidizing *Archaea* (AOA), and Ammonia Oxidizing *Bacteria* (AOB), and their response to physico-chemical gradients are not well known. To test if
30 different niches for potential nitrifiers exist in estuary systems, we assessed by pyrosequencing the diversity of archaeal gene transcripts marker for taxonomy (16SrRNA), and quantify by qPCR thaumarchaeal *amoA* and *ureC*, and bacterial *amoA* during an entire year along a salinity gradient in the surface zone of the Charente Estuary (Atlantic coast, France). This study provided compelling evidences of nitrifiers' successions from river to sea
35 with highest AOB activity in the freshwater station while AOA were predominant in the marine station. This clearly highlighted that AOA and AOB were adapted to different ecological niches associated with contrasting salinity. Interestingly, in the freshwater station, thaumarchaeal *ureC* transcripts abundance was the highest. Valuable insights into thaumarchaeal ecotypes through the phylogeny were also provided, and showed different
40 MGI subclusters associated with contrasted dynamics, with some able to degrade urea.

Introduction

The discovery of aerobic ammonia oxidation pathway within the domain of *Archaea* (Konneke et al., 2005; Treusch et al., 2005) has led to a dramatic shift in the classical view of
45 nitrification in ecosystems and, more generally, in the current model of nitrogen cycle. The wide distribution of Ammonia Oxidizing *Archaea* (AOA), affiliated with *Thaumarchaeota* Marine Group I (MGI), across a variety of aquatic environments is now well established through reports on the abundance of genes encoding archaeal ammonia monooxygenase (*amoA*) in oceanic waters (Francis *et al.*, 2005; Galand *et al.*, 2010; Karner *et al.*, 2001) and
50 freshwater ecosystems (Auguet *et al.*, 2011; Hugoni *et al.*, 2013a; Vissers *et al.*, 2013).

Nevertheless, in those ecosystems, relative contribution of AOA over Ammonia Oxidizing *Bacteria* (AOB) remains unclear and factors that regulate ammonia oxidizing microorganisms' activity and diversity have not yet been fully elucidated. While in marine ecosystems studies suggested that AOA outnumber AOB (Beman *et al.*, 2008; Mincer *et al.*, 55 2007; Wuchter *et al.*, 2006), some works conducted in freshwater ecosystems propose a more complex picture of nitrifiers' ecology (Hugoni *et al.*, 2013a). Additionally the discovery that some *Thaumarchaeota* may degrade urea as a nitrogen source and the absence of the *ureC* gene in the most successful MGI, *Nitrosopumilus maritimus*, raises new questions about the existence of different archaeal nitrifiers' ecotypes able to cope with different environmental 60 conditions or potential competitor (Alonso-Saez *et al.*, 2012; Walker *et al.*, 2010).

Microbial communities in estuaries and some coastal margins vary greatly in space and time because of sharp gradients in salinity and nutrients, among other properties (Herlemann *et al.*, 2011; Kirchman *et al.*, 2005). In estuaries, the mixing of freshwater and saltwater creates steep physico-chemical gradients that are accompanied by shifts in the 65 resident microbial communities (Crump and Baross, 2000). Community transitions from marine to freshwater environment are mainly due to salinity (Lozupone and Knight, 2007). Temperature, nitrite concentrations, and net primary productivity have however also been shown to produce major effects on the nitrifiers' community structure in estuaries and shifts in ammonia oxidizers have been frequently detected (Bernhard *et al.*, 2007; Mosier and 70 Francis, 2008; Santoro *et al.*, 2010). Archaeal diversity retrieved estuaries consisted usually in the combination of microorganisms related to both marine and freshwater but also to soils and sediments microorganisms (Crump and Baross, 2000). In riverine ecosystems the presence of both *Euryarchaeota* and *Thaumarchaeota* has been detected but their proportion varied according to the studied location. Indeed, while *Thaumarchaeota* dominated in the Rhine, 75 *Euryarchaeota* were the most abundant in the arctic Mackenzie River, suggesting that several

environmental parameters shaped archaeal diversity (Galand *et al.*, 2006; Herfort *et al.*, 2009). However, temporal surveys are lacking making difficult to address archaeal ecology. Moreover, the dynamics of aquatic microorganisms have traditionally been studied at the DNA level (Alonso-Saez *et al.*, 2007; Mary *et al.*, 2006), but recent studies have shown the importance of differentiating the active communities from the total communities by the use of both 16SrDNA and 16SrRNA (Campbell *et al.*, 2011; Hugoni *et al.*, 2013b; Lami *et al.*, 2009).

Despite the importance of *Archaea* in marine biogeochemical cycles (Schleper and Nicol, 2010; Walker *et al.*, 2010) little is known about the dynamics of archaeal populations and their activity across salinity and chemical gradient (Beman *et al.*, 2010; La Cono *et al.*, 2013; Vissers *et al.*, 2013). Here we used an estuary to test whether different niches exist for prokaryotic nitrifiers. Thaumarchaeal *amoA* and *ureC* transcripts were quantified in comparison to bacterial *amoA* and the community structure of active *Archaea* was characterized over one year along a salinity gradient in the surface layer of the Charente Estuary, on the west coast of France, by pyrosequencing 16S rRNA in the V3-V5 region.

Materials and methods

Study sites, sampling and chemical analyses

The Charente is a 350 km coastal river draining a 10,000 km² basin and emerging in the basin of Marennes Oleron. The sampling area (Figure 1) started at St-Savinien, upstream of the Charente (freshwater station) and ended in the Charente estuary (marine station), with one intermediary station (mesohaline station). Each station was characterized by a specific salinity class, ranging from 0 to 35PSU. Surface water (50 cm below the surface) was collected monthly in each station from April 2011 to March 2012, except May 2011 in the marine station, by using a 10-L Niskin Bottle (35 samples). Water temperature, salinity and dissolved

oxygen content were determined with a multiparameter probe (YSI GRANT 3800). Phosphorus (P-PO₄) and ammonia (N-NH₄) contents were analyzed using standard American Public Health Association (1992) methods. Chlorophyll *a* (Chl*a*) concentration was determined by spectrophotometry (Lorenzen, 1967b; Strickland and Parsons, 1968). The discharge data of the Charente at St-Savinien was obtained thanks to the Flood forecasting service (Conseil Général of Charente-Maritime).

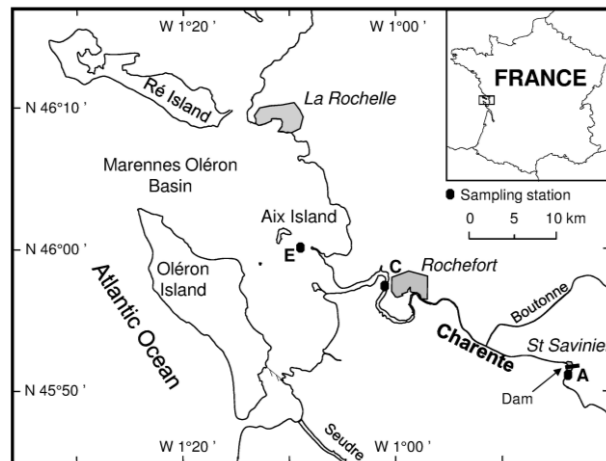


Figure 1. Location of sampling stations (median position over 12 sampling dates) along the Charente Estuary (map source: JC Auguet). The freshwater station was called A, the mesohaline station: C, and the marine station: E.

RNA extraction and pyrosequencing

A sub-sample (300 mL), added with an equal volume of RNA Later (ammonia sulfate 7.93M, sodium citrate 0.025M, EDTA 0.02M qsp 1.5L of RNase free water, pH 5.2), was pre-filtered through 5-μm pore-size polycarbonate filters (Millipore) and collected on 0.2-μm pore-size (pressure <10 kPa) polycarbonate filters (Millipore) and stored at -80°C until nucleic acid extraction. The RNA extraction method was modified from Hugoni *et al.* (2013b) using a combination of mechanic and enzymatic cell lysis, followed by extraction using the AllPrep DNA/RNA kit (Qiagen, Valencia, CA). The RNA samples were tested for the presence of contaminating genomic DNA by PCR and then reverse transcribed with random primers using the SuperScript[®] VILO (Invitrogen). The amplification of the V3-V5 region of

the 16S rRNA genes was performed with universal archaeal primers Arch349F (GYGCASCAGKCGMGAAG) and Arch806R (GGACTACVSGGGTATCTAAT) (Takai and Horikoshi, 2000), followed by pyrosequencing using a Roche 454 GS-FLX system with titanium chemistry by a commercial laboratory (MR.DNA, Shallowater, TX, USA).

Bioinformatic analysis

The 16S rRNA pyrosequencing dataset represented 175,637 raw sequences. All the sequences were checked for the following quality criteria: (i) no Ns; (ii) quality score ≥ 25 ; (iii) a minimum sequence length of 200 bp; and (iv) no sequencing error in the forward primer. The remaining reads were clustered at a 97% similarity threshold (Kim *et al.*, 2011b) and representative OTUs were inserted in phylogenetic trees for taxonomic annotation. The process was automated by PANAM that also computed richness and diversity indexes, Schao1 and Shannon respectively (<http://code.google.com/p/panam-phylogenetic-annotation/downloads/list>; (Taib et al., 2013)). After the removal of sequences affiliated with *Bacteria*, the dataset contained a total of 27,803 archaeal sequences distributed into 1825 OTUs (Supplementary Table 1).

Phylogenetic trees including OTUs retrieved in the 3 sampling stations were constructed by aligning reference sequences from the literature using Muscle (Edgar, 2004) and neighbor joining phylogenies were built using Mega5 (Tamura et al., 2011). This allowed us to determine major MGI clusters and evaluate the proportion of OTUs retrieved in each station.

RT quantitative PCR analysis

The qPCR protocol was modified from Hugoni *et al.* (2013a) on the cDNA produced from the 35 samples, using the primers described in the Table 1. Briefly, copy numbers of bacterial and

thaumarchaeal *amoA* in the environmental samples were determined in triplicate on the cDNA. The reaction mixture (25 µL) contained MESA GREEN qPCR MasterMix Plus for SYBR Assay[®] (1X, Eurogentec) added with 0.8 µg of BSA, 0.7 µM of primers and ultra-pure sterile water. One µL of cDNA was added to 24 µL of mix in each well. qPCR reactions consisted of an initial denaturing step at 94°C (for 15min for thaumarchaeal *amoA*, and 5min for thaumarchaeal *ureC* and bacterial *amoA* transcripts) followed by 40 cycles (thaumarchaeal *amoA*: 94°C 15sec, 52°C 30sec, 72°C 30sec; thaumarchaeal *ureC*: 94°C 1min, 55°C 1min, 72°C 2min and bacterial *amoA*: 95°C 30sec, 56°C 40sec, 72°C 2min). Standard curves were generated from mix of clone's representative of the environments studied. All reactions were performed with standard curves spanning from 10¹ to 10⁸ copies per µL. Mean PCR efficiencies and correlation coefficients for standard curves were as follows: for the thaumarchaeal *ureC* assay, 98 %, $r^2 = 0.98$, for the thaumarchaeal *amoA* assay, 108 %, $r^2 = 1.00$, and for the bacterial *amoA* assay, 107 %, $r^2 = 1.00$.

Table 1. Primers used for qPCR and pyrosequencing, and annealing conditions used in this study.

Application	Primer	Primer sequence 5' – 3'	Annealing temperature	Targeted group	Reference
qPCR	amoA-1F	GGGGTTTCTACTGGTGGT	56°C	β-proteobacterial <i>amoA</i>	Rotthauwe <i>et al.</i> , 1997
	AmoA-RNEW	CCCCTCBGSAAAVCCTTCTTC			
	CrenAmoAModF	TGGCTAAGACGMTGTA	52°C	Thaumarchaeal <i>amoA</i>	Mincer <i>et al.</i> , 2007
	CrenAmoAModR	AAGCGGCCATCCATCTGTA			
	Thaum-UreC F	ATGCAATYTGTAATGGAACWACWAC	55°C	Thaumarchaeal <i>ureC</i>	Alonso-Saez <i>et al.</i> , 2012
	Thaum-UreC R	AGTTGTYCCCAATCTTCATGTAATTTTA	55°C	Thaumarchaeal <i>ureC</i>	
Pyrosequencing	Arch349F	GYGCASCAGKCGMGAAW	55°C	Archaeal 16S rRNA	Takai <i>et al.</i> , 2000
	Arch 806R	GGACTACVSGGTATCTAAT			

Statistical analysis

Statistical analyses were conducted using R associated to the package VEGAN (<http://cran.r-project.org/web/packages/vegan/index.html>).

Canonical correspondence analysis (CCA) was performed to assess the relationships between active archaeal taxonomic groups and environmental parameters. CCA was performed on 6 environmental factors (temperature, salinity, Chl*a* content, pH, phosphate, and ammonium concentrations) and the taxonomic groups abundance matrix (inferred from 16SrRNA reads number).

To explain the variation of archaeal *amoA* and *ureC* transcripts abundance, redundancy analysis (RDA) was used after a forward selection (Borcard *et al.*, 1992) of the ten thaumarchaeal OTUs susceptible to explain a significant part of changes in archaeal *amoA* and *ureC* transcripts abundance (inferred from the qPCR assays).

Results

Physico-chemical and biological characteristics of the Charente estuary

The Charente Estuary was investigated through three stations along a salinity gradient sampled monthly during one year (Supplementary Table 1). The freshwater, mesohaline and marine stations were characterized by a mean salinity of 0, 14.9 and 33.2 PSU and a mean Chl*a* of 17.6, 4.34 and 2.74 $\mu\text{g.L}^{-1}$ respectively. Ammonia concentrations were on average lower in the marine station than in the freshwater (0.019 mg.L^{-1} and 0.081 mg.L^{-1} respectively) and conversely for phosphate concentrations (on average 0.096 mg.L^{-1} and 0.038 mg.L^{-1} respectively). The river discharge ranged between 36.2 and 29.2 $\text{m}^3.\text{sec}^{-1}$ from April to October 2011, while it increased from November 2011 (64.9 $\text{m}^3.\text{sec}^{-1}$) to January 2012 (142.1 $\text{m}^3.\text{sec}^{-1}$) and decreased in February and March 2012 (63.3 and 36.3 $\text{m}^3.\text{sec}^{-1}$ respectively).

Active ammonia oxidizing and urea degrading communities' dynamics

To better understand the role of both *Archaea* and *Bacteria* in ammonia oxidation, the transcripts quantification of their *amoA* genes were examined by quantitative PCR during a one-year period (Figure 2.A and B). Few archaeal *amoA* transcripts were detected in the freshwater station (from 0 to 13 copies of transcripts.mL⁻¹), nevertheless three distinct periods of activity could be observed: one from April to September, then a second from October to November and finally from December to March. The mesohaline station was also characterized by a AOA activity from April to September and then, in October and November, the greatest peak being observed during June and July (about 115-127 copies of transcripts.mL⁻¹) (Figure 2.A). Conversely, the more important AOA activity at the marine station was observed only from October to January (from 1 to 91 copies), the rest of the year the copy number of transcripts per mL ranging from 1 to 10. In contrast, bacterial *amoA* transcripts level peaked twice in the year at the freshwater station, first in May and June and then in from August to November, reaching up to 225 times more transcripts than for AOA (Figure 2.B). In the mesohaline station, AOB transcripts showed overall a similar dynamic to that of AOA, although their number was generally lower. In the marine station, bacterial *amoA* transcription increase took place earlier than that of AOA, being detected between from August to October.

To assess the genetic potential of *Thaumarchaeota* for ureolytic nitrogen metabolism in these environments, *ureC* transcripts abundance was also determined using qPCR (Figure 2.C) during a one-year period. In the Charente Estuary, the freshwater station presented the most important transcripts abundance (maximum of 10 copies of transcripts.mL⁻¹) and maximal abundance was associated with the same three periods of archaeal *amoA* activity. In both the mesohaline and marine stations, a low level of *ureC* transcripts was detected all the

year excepted in May and June at the mesohaline station, which was congruent with the highest archaeal *amoA* genes expression.

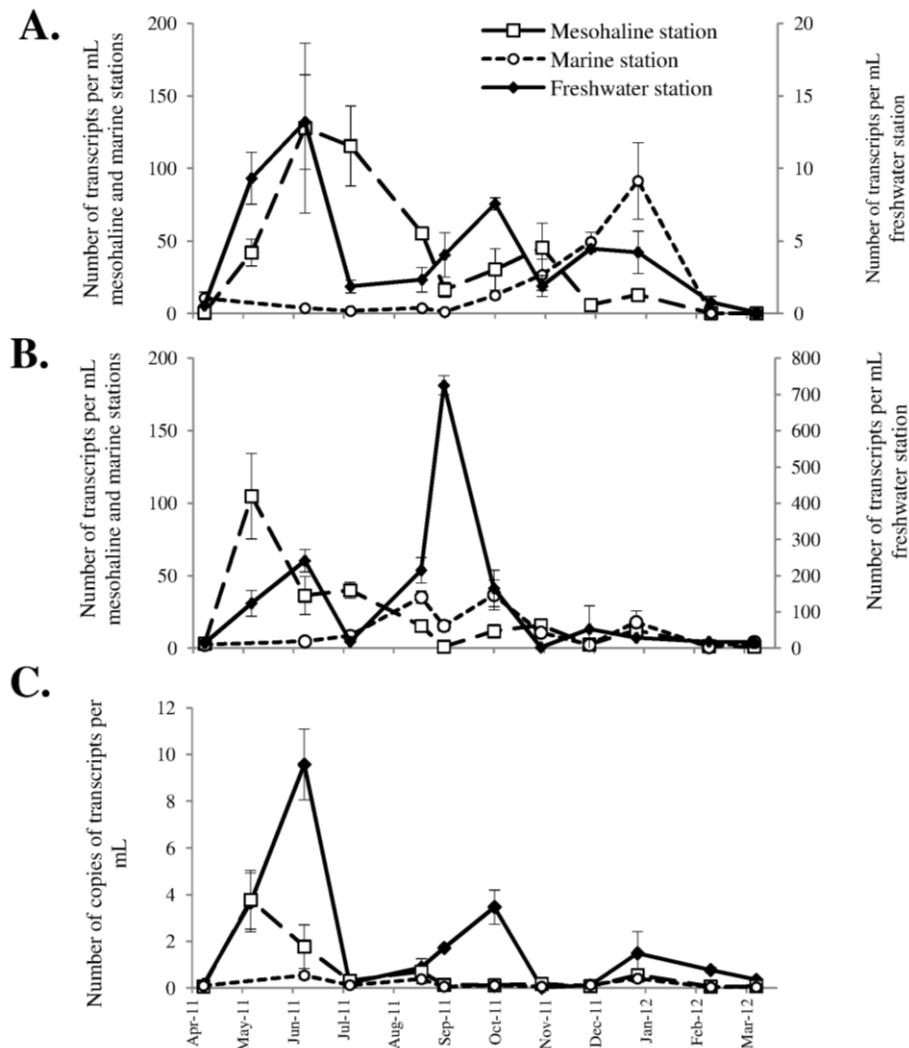


Figure 2. Abundances of thaumarchaeal (A) and bacterial (B) *amoA* transcripts and (C) thaumarchaeal *ureC* transcripts per mL during the one year survey.

Active *Archaea* community structure and dynamics

The changes in the community structure of active archaeal populations were evaluated over time from river to sea by using pyrosequencing. In the mesohaline station, 923 OTUs were observed against 550 and 352 in the freshwater and marine stations respectively. Overall the diversity was highest in the freshwater and lowest in the marine station (Supplementary Figure 1).

We evaluated active archaeal diversity on the basis of the three main periods of AOA activity defined above. Pyrosequencing did not work on marine samples (no amplification) and thus we could not define archaeal diversity from October to November in this station. In the freshwater station, different patterns of community composition were visible with time (Figure 3). The active MGI sequences were more abundant from April to November representing about 83% of the sequences. These *Thaumarchaeota* were dominated by 11 OTUs with best match by Blast to sequences recovered from lacustrine freshwaters, groundwater, rivers and mangrove sediments. In contrast, the methanogenic lineages that represented less than 10% of the sequences from April to November, and increased from December to March reaching up to 60% of the sequences. Five methanogenic OTUs belonging to the *Methanosaeta* genus dominated. They were closely related to sequences found in salt marsh and lacustrine sediments, but also retrieved in groundwater and wetlands. In the mesohaline station, the active archaeal community was clearly dominated by *Thaumarchaeota* MGI all year round (between 86% and 93% of the sequences, Figure 3). For *Euryarchaeota*, few methanogenic lineages were retrieved in the mesohaline station and, MGII group represented about 7% of the sequences between December and March.

In the marine station, we could also distinguish different archaeal seasons. From April to September, *Thaumarchaeota* MGI represented about 17% of the 16SrRNA transcripts, the remaining being near exclusively *Euryarchaeota* MGII sequences (around 81%). Conversely, from December to March, thaumarchaeal MGI represented about 70% of the sequences while euryarchaeal MGII sequences accounted for 27% of the sequences. Six abundant thaumarchaeal OTUs were related to lacustrine freshwater, groundwater, rivers and mangrove sediment but also to ocean waters sequences. On the other hand, the fifteen euryarchaeal abundant OTUs affiliated were related to marine and oceanic water sequences.

The canonical correspondence analysis plot showed that MGII was mainly related to salinity and consequently to the marine station. MGI activity was linked to the mesohaline station, while the majority of euryarchaeal groups (i.e. LDS, RC-V, MEG) were related to ammonia and to the freshwater station (Supplementary Figure 2).

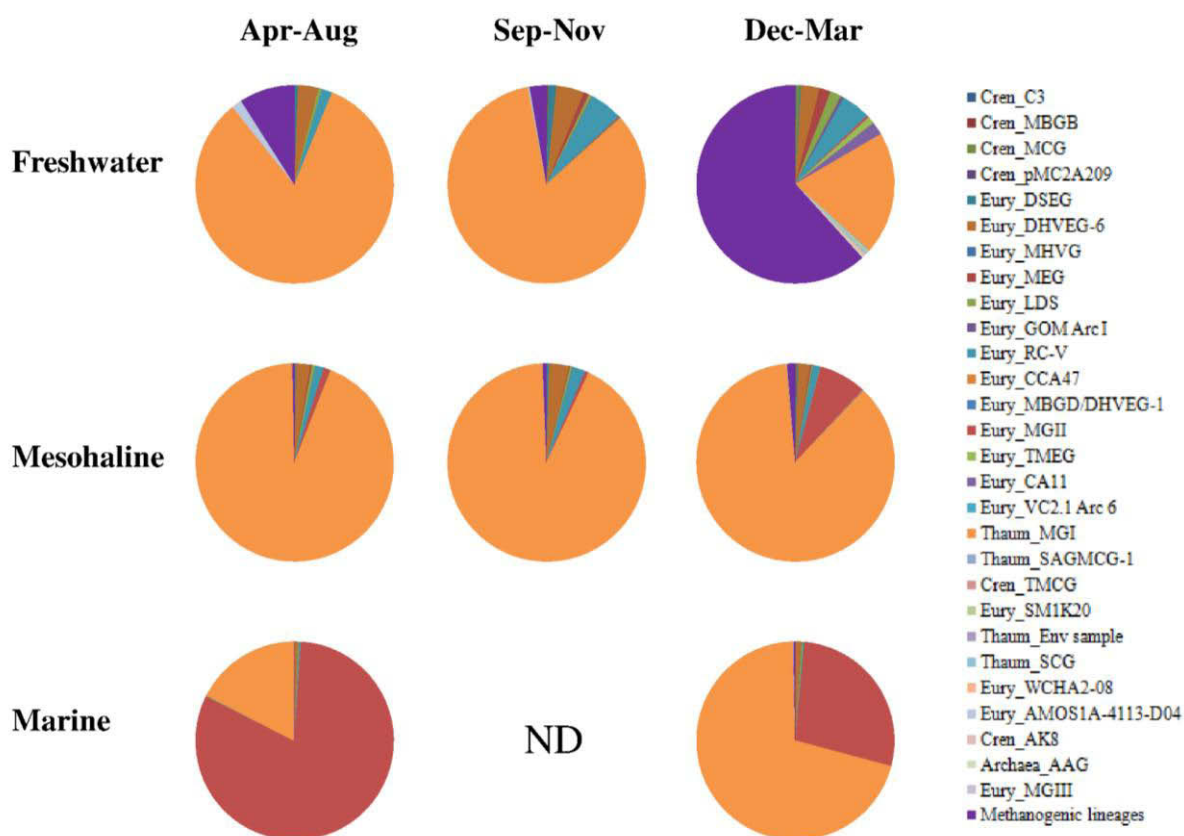
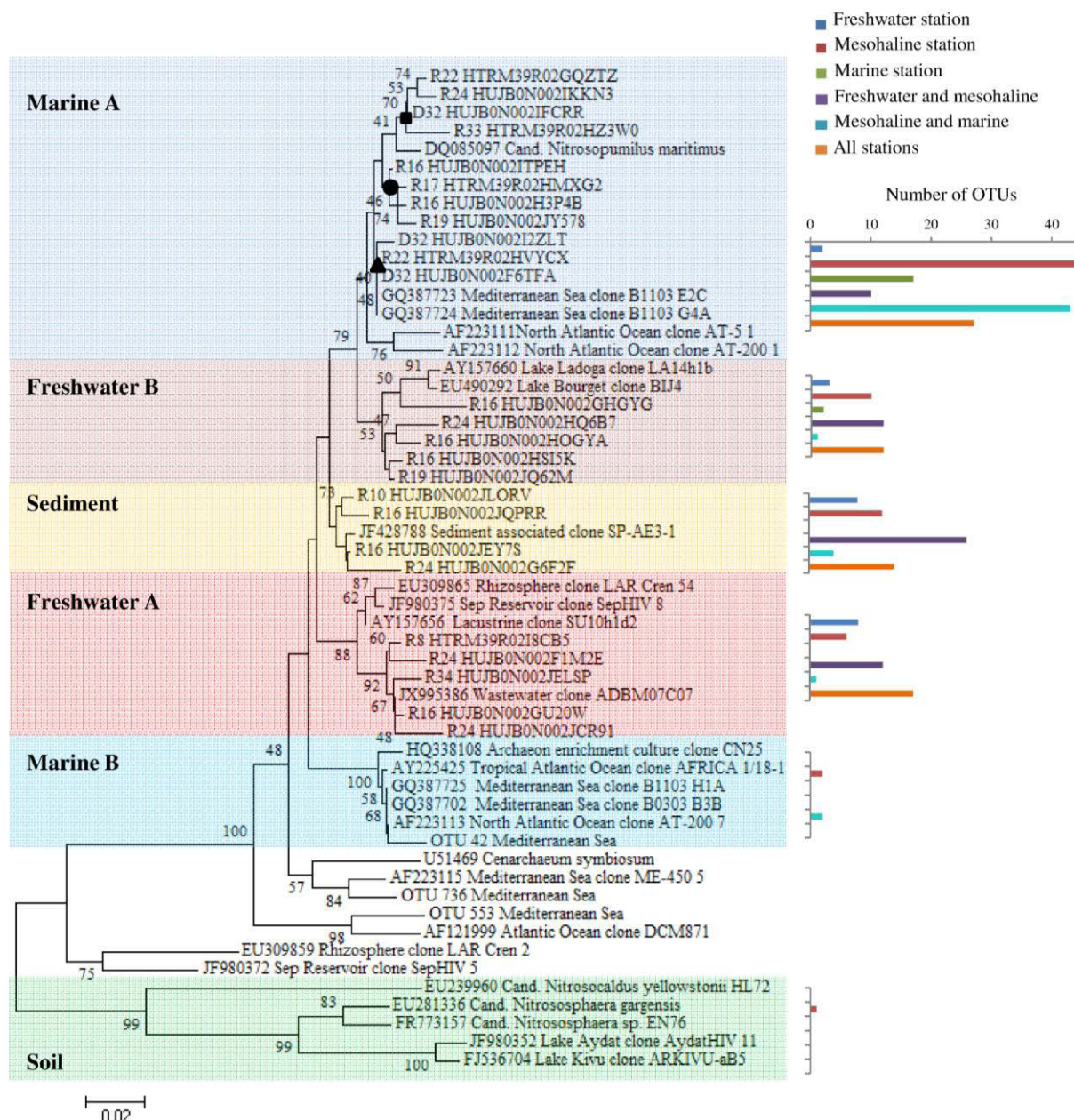


Figure 3. Sequences proportion of 16S rDNA transcripts for taxonomic groups retrieved in the freshwater station, the mesohaline station, and the marine station during three distinct periods: from April to September, from October to November and from December to March.

Thaumarchaeal diversity

Active thaumarchaeal OTUs retrieved in the 3 sampling stations fell into six major MGI subclusters (Figure 4). Among them, Marine A and B were initially retrieved in marine ecosystems (Hugoni et al., 2013b) while Freshwater A and B, in freshwater ones (Hugoni et al., 2013a). Interestingly, we also found OTUs belonging to a subcluster related to sedimentary ecosystems. OTUs present in the freshwater station belonged mainly to the Freshwater A and Sediment associated groups, while those in the mesohaline and the marine

270 stations clustered mainly with sequences belonging to the Marine A subcluster. Common
OTUs from the freshwater and the mesohaline stations fell in the Sediment subcluster while
those that were characteristic of the mesohaline and marine station mostly fell in the Marine
A subcluster.



275 **Figure 4.** Phylogenetic tree including active *Thaumarchaeota* MGI OTUs retrieved in the three
sampling stations (freshwater, mesohaline and marine) of the Charente estuary. Bootstrap values >40
are shown expressed as a percentage of 500 replicates. Histograms on the right represented the number
of OTUs retrieved in each station and in the freshwater and mesohaline, the mesohaline and marine
and in all stations. Seasonal changes in thaumarchaeal OTUs from the Marine A cluster were
280 illustrated for OTUs in the mesohaline station which were active from October to November (■), from
April to September (●), and from December to March (▲).

Linking thaumarchaeal MGI diversity and nitrogen metabolism

We focused on the abundant *Thaumarchaeota* OTUs retrieved in each sampling station, and observed contrasted activity periods (i.e. 16SrRNA transcripts) suggesting different niches for

285 *Thaumarchaeota* ecotypes.

First, in the freshwater station, when AOB were active, thaumarchaeal *amoA* transcripts were less abundant, and thaumarchaeal *ureC* transcripts could be detected following the same dynamic as *amoA* AOA. We therefore investigated abundant MGI OTUs dynamics in this freshwater station and observed that there were more MGI sequences in
290 October and November (Supplementary Table 2). At the same period, thaumarchaeal *amoA* and *ureC* activities were detected.

In contrast, in the mesohaline station, abundant MGI OTUs were mainly active during the April to September period. Nevertheless, temporal dynamics of abundant MGI OTUs affiliated with the Marine A subcluster showed major changes in composition along the
295 season. Indeed, we could distinguish 3 subclusters in the MGI phylogeny (Figure 4) associated with contrasted seasonal dynamics (Supplementary Table 2): one was active from April to September and represented by OTUs R16_HUJB0N002H3P4B, R19_HUJB0N002JY578, R17_HTRM39R02HMXG2 and R16_HUJB0N002ITPEH, the second in October and November (OTUs D32_HUJB0N002IFCRR and
300 R22_HTRM39R02GQZTZ) and the last one, in December and March (OTUs R22_HTRM39R02HVYCX, D32_HUJB0N002F6TFA, and D32_HUJB0N002I2ZLT). In this station, *amoA* AOA transcripts abundance outnumbered bacterial ones and three different periods were associated with *amoA* AOA and thaumarchaeal *ureC* activity. Surprisingly, those three periods were linked with the Marine A subclusters highest activity, suggesting
305 distinct MGI OTUs that could be active successively during the year, associated with

contrasted environmental parameters. Thus, we hypothesized that Marine A subclusters were ecotypes.

In the marine station the lack of information from October to November did not allow to conclude. Nevertheless, within the cluster Marine A, OTUs R22_HTRM39R02HVCYCX, D32_HUJB0N002F6TFA, and D32_HUJB0N002I2ZLT were more active from December to March (Supplementary Table 2). This was congruent with the highest *amoA* transcripts abundance retrieved at the same period in this marine station and suggested that those particular OTUs were more active in marine conditions than in mesohaline ones.

To gain insight into 16S rRNA OTUs and functional genes, we used common thaumarchaeal abundant OTUs retrieved in all sampling stations in a forward RDA to explain quantitative variations of *amoA* and *ureC* transcripts across time. Then, the ten best informative OTUs (i.e. accounting for 77% of the total variation) were selected to build a RDA. The RDA plot showed that the ureolytic metabolism could be linked to three OTUs (R16_HUJB0N002GU20W, R24_HUJB0N002JCR91 and R10_HUJB0N002JLORV). Those OTUs were affiliated with the Sediment and Freshwater A subclusters. On the other hand, the *amoA* potential was correlated with six OTUs among which five were affiliated with the Marine A subcluster while the last one was affiliated with the Freshwater B subcluster (i.e. R17_HTRM39R02HMXG2, R16_HUJB0N002ITPEH, R22_HTRM39R02HVCYCX, D32_HUJB0N002F6TFA, D32_HUJB0N002I2ZLT and R8_HTRM39R02IN9W4).

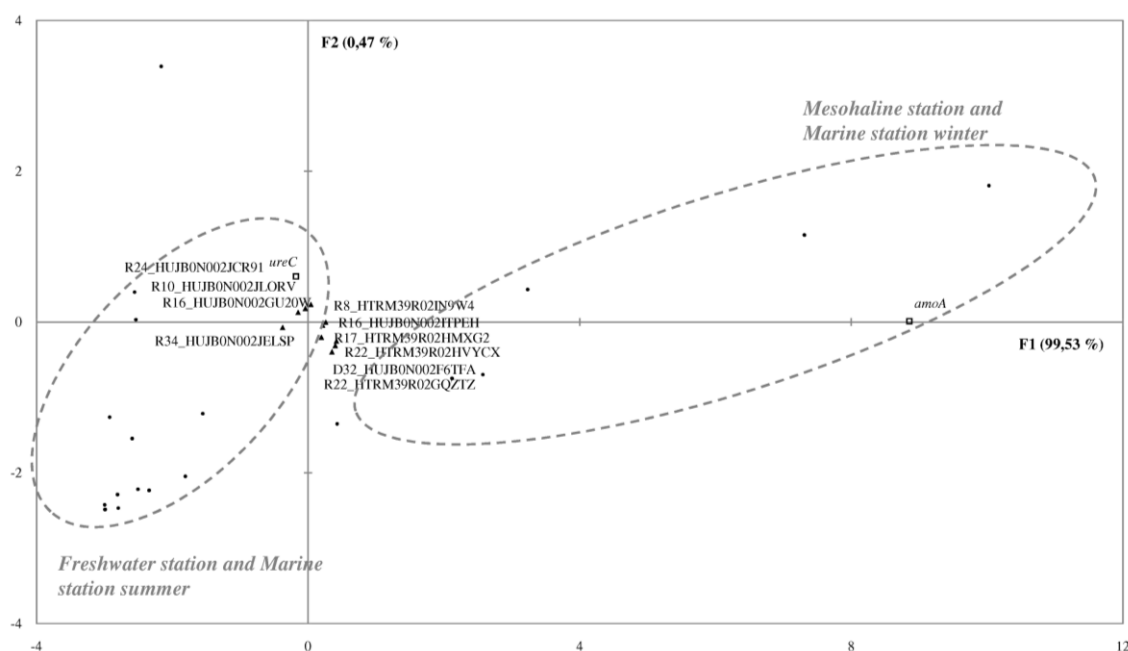


Figure 5. RDA plot of *ureC* and *amoA* transcripts abundance compared with active abundant thaumarchaeal OTUs.

Discussion

The quantification of ammonia oxidizing communities' *amoA* transcripts abundances in the water column of the Charente estuary over a 1-year period highlighted a transition in dominant ammonia-oxidizing populations between freshwater and marine stations and suggested that AOB activity decreased as salinity increased. The retrieval of higher AOB transcripts number in the freshwater station is congruent with previous studies suggesting that AOB potential nitrification rates were inhibited in high salinity environment (around 30 PSU (Bernhard et al., 2007)). In contrast, AOA activity was very low in these freshwater conditions while intermediary conditions (mesohaline station) favored their activity through the whole year. Surprisingly, some previous studies, DNA based, showed that AOA were more abundant in freshwater zones (Abell *et al.*, 2010; Mosier and Francis, 2008; Santoro *et al.*, 2008), nevertheless, if the relative distributions of AOA and AOB were considered in relation to ammonia concentrations rather than salinity, agreement between this and previous

studies was found. Indeed, maximal thaumarchaeal transcripts were retrieved in both mesohaline and marine stations when ammonia concentrations were the lowest. In the marine station, archaeal *amoA* activity was most important during the winter period, as previously suggested for *amoA* archaeal abundance in marine ecosystems (Galand et al., 2010). It has been suggested that the winter period was favorable for *Thaumarchaeota* to compete for ammonia compared to *Bacteria* (Winter et al., 2010). Nevertheless, in our study, AOA activity could not be statistically linked to any environmental parameters, suggesting other factors (i.e. physico-chemical parameters not considered in our work or top down controls) shaping AOA activity in this ecosystem.

Among MGI, we could define the presence of ecotypes presenting seasonal activity shifts and associated with potential ammonia oxidation activity. Indeed, typical clusters from freshwater and marine environments were retrieved and major changes in composition were observed among the Marine A subcluster, suggesting a phylogenetic diversity associated with contrasted seasons. The variable activity levels for the different MGI subclusters could suggest an adaptation to different niches and may indicate an ecological specialization, comforting the idea of separate ecotypes, as previously shown for *Vibrionaceae* populations (Szabo et al., 2013). Although we could not precisely define ecological niches in the present study, we hypothesized that the MGI subclusters represented ecotypes, as previously demonstrated for *Bacteria* (Coleman and Chisholm, 2010; Schwalbach et al., 2010). It has been reported that *Thaumarchaeota* could use urea to fuel nitrification and could thus represent an alternative metabolic pathway when the availability of ammonia was limited (Alonso-Saez et al., 2012). In our work, we assessed that specific adaptations of MGI to peculiar niches could come with the use of ureolysis to supply the ammonia oxidation when AOB activity was predominant. This was illustrated by the retrieval of *ureC* transcripts in the freshwater station, when AOB dominated for the ammonia oxidation process. Moreover, in

our study this ability was linked to some MGI thaumarchaeal OTUs affiliated with the Freshwater A and Sediment subclusters and it seemed that not all MGI OTUs might be able to
370 degrade urea, as illustrated by the weak *ureC* transcripts number retrieved in the marine station. This could also be explained through the thaumarchaeal *amoA* phylogeny revealing that all OTUs were related to the *Nitrosopumilus* cluster, represented by *N. maritimus* which does not possess genes for urea utilization (Walker *et al.*, 2010). This study clearly highlighted that some further work are necessary to understand which parameters shaped
375 AOA involvement into ammonia or urea utilization. Nevertheless, we could not assume that *amoA*-carrying *Archaea* are oxidizing ammonia and that transcripts detection was not sufficient to affirm ammonia oxidation involvement (Pester *et al.*, 2011).

The co-occurrence of both MGI and methanogenic groups in the freshwater station confirms the hypothesis that methanogenic lineages thrive in freshwater sediments and water
380 column rather than in marine and brackish ones (Carbonero *et al.*, 2012). The retrieval of some active methanogens in the mesohaline station could suggest that they were slowly washed away from their habitat (i.e. freshwater), resulting in decreasing activity with increasing salinity (Pulliam, 1988). Some typical freshwater groups like LDS or RC-V clusters were also retrieved (4.8 and 11.9 % of the reads respectively) in this non saline
385 station suggesting an eventual drain from nearby rivers. Indeed, the RC-V cluster was first detected in anoxic rice paddy sediments (Großkopf *et al.*, 1998) while LDS cluster was first identified in Lake Dagow Sediment (Glissman *et al.*, 2004). Those two groups were also identified in rivers (Galand *et al.*, 2006; Herfort *et al.*, 2009) where they could account for a large proportion of the archaeal cell counts (Herfort *et al.*, 2009). In the marine station, there
390 was a clear succession between MGI and MGII, with a predominance of active MGI during the winter period, as previously shown in the North-Western Mediterranean Sea (Hugoni *et al.*, 2013b), while MGII dominated the active archaeal assemblage during the summer period.

The ecological success of this group in the marine station could come from the putative presence of genes encoding the proteorhodopsin (*pop*), already retrieved in members of the *Euryarchaeota* MGII.A in the surface waters of the North Pacific (Frigaard et al., 2006) and in the reconstruction of a coastal MGII.A genome (Iverson et al., 2012). We therefore hypothesized that phototrophic metabolism could be present in the Charente estuary when MGII could use light as an energy source, explaining their summer activity, as irradiance is more important in this season.

Our study confirms the activity of some minors groups, like Crenarchaeota Group C3, Miscellaneous Crenarchaeotic Group (MCG) and Marine Benthic Group B (MBG-B) retrieved in the water column of station A and C. This suggested that those groups considered as ubiquitous (Inagaki et al., 2006) and usually retrieved in estuarine sediments (Singh et al., 2010) could be active in estuaries water column. This was supported by a recent work conducted through a single-cell approach that proposed a role of MCG in protein remineralization in anoxic sediments (Lloyd et al., 2013).

In summary, this study clearly showed that nitrification plays an important role in estuarine biogeochemistry, yet efforts to link this process to the microorganisms that mediate it were surprisingly limited. We assessed a niche partitioning of active prokaryotic nitrifiers correlated with salinity. Among thaumarchaeal nitrifiers, different niches were observed among specific subclusters assessing a microdiversity linked with seasonal changes. Ammonia oxidation was the first and rate-limiting step of nitrification, yet alternative pathways to fuel ammonia oxidation were recently described and could impact relative contributions of *Bacteria* and *Archaea* in this process. Previous large-scale studies of AOA *amoA* (Francis et al., 2005) and the archaeal 16S rRNA (Auguet et al., 2009) have recorded close evolutionary relatedness between organisms inhabiting physico-chemically similar environments, independent of geographical distance between the environment. This suggested

that the prevailing geochemistry, rather than localized dispersal, is the major driving factor determining archaeal OTU distribution (Bouskill et al., 2012), and raised many questions about the niche conservatism principle (Wiens *et al.*, 2010). Nevertheless understanding environmental parameters that affect active ecotypes distribution is still a challenge for microbial ecologists.

Acknowledgments

We thank P. Pineau, N. Lachaussée, M. Breret, F. Mornet, L. Beaugéard, J. Lavaud and J. Jourde for samples. We thank A. Vellet and I. Louati for their technical support during the experimentations, F. Toubanc, T. Coulombier and I. Brenon for active discussion on the output of the Charente estuary, J.C. Auguet for providing us the map of the sampling location of the sampling stations. This work was supported by a CNRS Program Ecosphère Continentale et Côtière (EC2CO, 2010–2012).

Additional Supporting Information may be found in the online version of this article.

References

- Abell GC, Revill AT, Smith C, Bissett AP, Volkman JK, Robert SS (2010). Archaeal ammonia oxidizers and nirS-type denitrifiers dominate sediment nitrifying and denitrifying populations in a subtropical macrotidal estuary. *ISME J* **4**: 286-300.
- Alonso-Saez L, Balague V, Sa EL *et al* (2007). Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiol Ecol* **60**: 98-112.
- Alonso-Saez L, Waller AS, Mende DR *et al* (2012). Role for urea in nitrification by polar marine *Archaea*. *Proc Natl Acad Sci USA* **109**: 17989-94.
- Auguet JC, Barberan A, Casamayor EO (2009). Global ecological patterns in uncultured *Archaea*. *ISME J* **4**: 182-90.
- Auguet JC, Nomokonova N, Camarero L, Casamayor EO (2011). Seasonal changes of freshwater ammonia-oxidizing archaeal assemblages and nitrogen species in oligotrophic alpine lakes. *Appl Environ Microbiol* **77**: 1937-45.
- Beman JM, Popp BN, Francis CA (2008). Molecular and biogeochemical evidence for ammonia oxidation by marine *Crenarchaeota* in the Gulf of California. *ISME J* **2**: 429-41.
- Beman JM, Sachdeva R, Fuhrman JA (2010). Population ecology of nitrifying *archaea* and *bacteria* in the Southern California Bight. *Environ Microbiol* **12**: 1282-92.
- Bernhard AE, Tucker J, Giblin AE, Stahl DA (2007). Functionally distinct communities of ammonia-oxidizing bacteria along an estuarine salinity gradient. *Environ Microbiol* **9**: 1439-47.
- Borcard D, Legendre P, Drapeau P (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045-1055.
- Bouskill NJ, Eveillard D, Chien D, Jayakumar A, Ward BB (2012). Environmental factors determining ammonia-oxidizing organism distribution and diversity in marine environments. *Environ Microbiol* **14**: 714-29.
- Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011). Activity of abundant and rare *bacteria* in a coastal ocean. *Proc Natl Acad Sci USA* **108**: 12776-81.
- Carbonero F, Oakley BB, Hawkins RJ, Purdy KJ (2012). Genotypic distribution of a specialist model microorganism, *Methanosaeta*, along an estuarine gradient: does metabolic restriction limit niche differentiation potential? *Microb Ecol* **63**: 856-64.
- Coleman ML, Chisholm SW (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634-9.
- Crump BC, Baross JA (2000). Archaeaplankton in the Columbia River, its estuary and the adjacent coastal ocean, USA. *FEMS Microbiol Ecol* **31**: 231-239.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-7.

495 Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005). Ubiquity and diversity of ammonia-oxidizing *archaea* in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* **102**: 14683-8.

500 Frigaard NU, Martinez A, Mincer TJ, DeLong EF (2006). Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and *Archaea*. *Nature* **439**: 847-50.

Galand PE, Gutierrez-Provecho C, Massana R, Gasol J, Casamayor EO (2010). Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea. *Limnol Oceanogr* **55**: 2117-2125.

505 Galand PE, Lovejoy C, Vincent WF (2006). Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquatic Microbial Ecology* **44**: 115-126.

510 Glissman K, Chin KJ, Casper P, Conrad R (2004). Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature. *Microb Ecol* **48**: 389-99.

515 Großkopf R, Stubner S, Liesack W (1998). Novel Euryarchaeotal Lineages Detected on Rice Roots and in the Anoxic Bulk Soil of Flooded Rice Microcosms. *Appl Environ Microbiol* **64**: 4983-4989.

520 Herfort L, Kim JH, Coolen MJL *et al* (2009). Diversity of *Archaea* and detection of crenarchaeotal *amoA* genes in the river Rhine and Têt. *Aquatic Microbial Ecology* **55**: 189-201.

Herlemann DP, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ, Andersson AF (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* **5**: 1571-9.

525 Hugoni M, Etien S, Bourges A *et al* (2013a). Dynamics of ammonia-oxidizing *Archaea* and *Bacteria* in contrasted freshwater ecosystems. *Res Microbiol* **164**: 360-70.

530 Hugoni M, Taib N, Debroas D *et al* (2013b). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci USA* **110**: 6004-9.

535 Inagaki F, Nunoura T, Nakagawa S *et al* (2006). Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proc Natl Acad Sci USA* **103**: 2815-20.

Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV (2012). Untangling genomes from metagenomes: revealing an uncultured class of marine *Euryarchaeota*. *Science* **335**: 587-90.

540

Karner MB, DeLong EF, Karl DM (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-10.

Kim M, Morrison M, Yu Z (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* **84**: 81-7.

Kirchman DL, Dittel AI, Malmstrom RR, Cottrell MT (2005). Biogeography of major bacterial groups in the Delaware Estuary. *Limnol Oceanogr* **50**: 1697-1706.

Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543-6.

La Cono V, La Spada G, Arcadi E *et al* (2013). Partaking of *Archaea* to biogeochemical cycling in oxygen-deficient zones of meromictic saline Lake Faro (Messina, Italy). *Environ Microbiol* **15**: 1717-33.

Lami R, Ghiglione JF, Desdevises JF, West NJ, Lebaron P (2009). Annual patterns of presence and activity of marine bacteria monitored by 16S rDNA-16SrRNA fingerprints in the coastal NW Mediterranean Sea. *Aquat Microb Ecol* **54**: 199-210.

Lloyd KG, Schreiber L, Petersen DG *et al* (2013). Predominant *archaea* in marine sediments degrade detrital proteins. *Nature* **496**: 215-8.

Lorenzen CJ (1967). Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnol. Oceanogr.* **12**: 343-346.

Lozupone CA, Knight R (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104**: 11436-40.

Mary I, Cummings DG, Biegala IC, Burkill PH, Archer SD, Zubkov MV (2006). Seasonal dynamics of bacterioplankton community structure at a coastal station in the western English Channel *Aqua Microbial Ecol* **42**: 119-126.

Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF (2007). Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ Microbiol* **9**: 1162-75.

Mosier AC, Francis CA (2008). Relative abundance and diversity of ammonia-oxidizing *archaea* and bacteria in the San Francisco Bay estuary. *Environ Microbiol* **10**: 3002-16.

Pester M, Schleper C, Wagner M (2011). The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol* **14**: 300-6.

Pulliam HR (1988). Sources, sinks, and population regulation. *The American Naturalist* **132**: 652-661.

Santoro AE, Casciotti K, Francis CA (2010). Activity, abundance and diversity of nitrifying archaea and bacteria in the central California Current. *Environmental Microbiology Reports* **12**: 1989-2006.

Santoro AE, Francis CA, de Sieyes NR, Boehm AB (2008). Shifts in the relative abundance of ammonia-oxidizing *bacteria* and *archaea* across physicochemical gradients in a subterranean estuary. *Environ Microbiol* **10**: 1068-79.

595 Schleper C, Nicol GW (2010). Ammonia-oxidising archaea--physiology, ecology and evolution. *Adv Microb Physiol* **57**: 1-41.

Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ (2010). The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* **12**: 490-500.

Singh SK, Verma P, Ramaiah N, Chandrashekar AA, Shouche YS (2010). Phylogenetic diversity of archaeal 16S rRNA and ammonia monooxygenase genes from tropical estuarine sediments on the central west coast of India. *Res Microbiol* **161**: 177-86.

605 Strickland J, Parsons T (1968). *A practical handbook of sea water analysis*.

Szabo G, Preheim SP, Kauffman KM *et al* (2013). Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J* **7**: 509-519.

610 Taib N, Mangot JF, Domaizon I, Bronner G, Debroas D (2013). Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS One* **8**: e58950.

615 Takai K, Horikoshi K (2000). Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Appl Environ Microbiol* **66**: 5066-72.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731-9.

625 Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic *crenarchaeota* in nitrogen cycling. *Environ Microbiol* **7**: 1985-95.

Vissers EW, Anselmetti FS, Bodelier PL *et al* (2013). Temporal and spatial coexistence of archaeal and bacterial amoA genes and gene transcripts in Lake Lucerne. *Archaea* **2013**: 289478.

630 Walker CB, de la Torre JR, Klotz MG *et al* (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107**: 8818-23.

635 Wiens JJ, Ackerly DD, Allen AP *et al* (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* **13**: 1310-24.

640 Winter C, Bouvier T, Weinbauer MG, Thingstad TF (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiol Mol Biol Rev* **74**: 42-57.

Wuchter C, Abbas B, Coolen MJ *et al* (2006). Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* **103**: 12317-22.

645

650

655

660

665

670

675

680

685

690

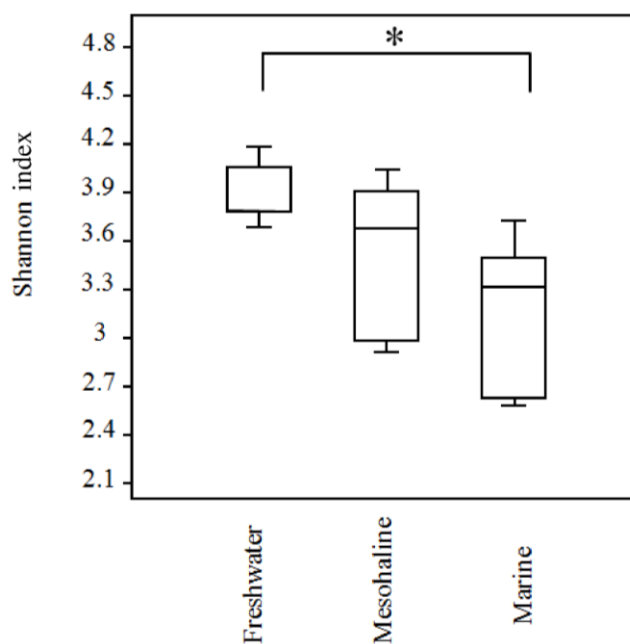
Supplementary Data

Supplementary Table 1. QC and *Archaea* affiliated sequences obtained for each sample from surface water column collected monthly in the Charente estuary. Environmental parameters (temperature, salinity, pH and Chl*a*, ammonia and phosphates concentrations) associated to each point are presented. QC, quality checked; ND, not determined.

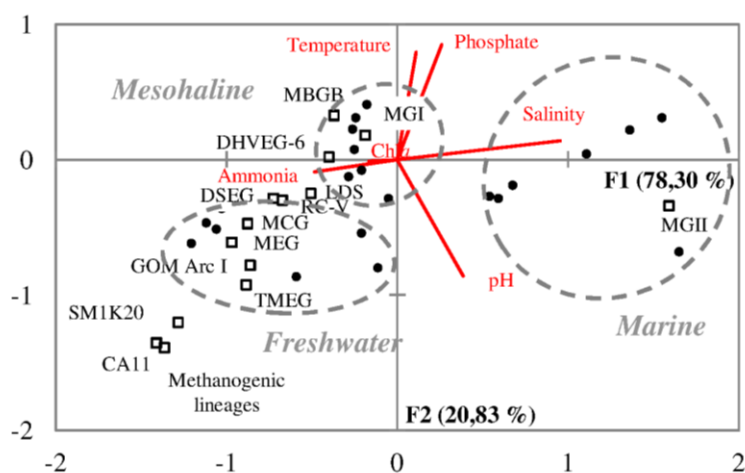
	Date	QC seq.	Archaea affiliated sequences	Temp. (°C)	Sal. (PSU)	pH	Chla (µg.L ⁻¹)	NH ₄ ⁺ (mg.L ⁻¹)	PO ₄ ³⁻ (mg.L ⁻¹)
		16SrRNA							
Freshwater	14/04/2011	ND	ND	16.9	0	8.65	3.618	0.113	0.008
	13/05/2011	ND	ND	20.7	0	8.39	6.686	0.030	0.018
	15/06/2011	2966	177	21.7	0	8.28	28.48	0.081	0.008
	13/07/2011	4163	52	23.7	0	8.04	5.658	0.143	0.082
	26/08/2011	1613	20	23.5	0	8.18	63.665	0.036	0.111
	09/09/2011	3351	116	21.7	0	8.13	56.069	0.049	0.021
	10/10/2011	2057	212	18.3	0	8.22	33.94	0.065	0.056
	08/11/2011	3055	1616	13.8	0	8.14	2.031	0.078	0.034
	08/12/2011	4909	394	11.5	0	8.21	0.691	0.122	0.014
	06/01/2012	5171	592	10.5	0	8.14	1.57	0.068	0.024
	20/02/2012	4501	946	7.2	0	8.1	7.089	0.081	0.011
	19/03/2012	306	35	12.1	0	ND	1.751	0.108	0.069
Mesohaline	14/04/2011	ND	ND	15.4	14.3	8.48	5.264	0.0	0.011
	13/05/2011	ND	ND	19.4	14.7	8.5	3.296	0.011	0.011
	15/06/2011	734	356	21	14.5	7.91	3.873	0.317	0.043
	13/07/2011	16263	7229	22.1	15.4	7.78	5.054	0.043	0.114
	26/08/2011	1610	797	22.8	14.7	7.89	9.805	0.064	0.079
	09/09/2011	ND	ND	21.1	14.3	7.58	4.276	0.0	0.069
	10/10/2011	4136	2798	18	14.7	7.79	4.669	0.032	0.085
	08/11/2011	1169	724	13.7	14.3	7.9	2.736	0.033	0.056
	08/12/2011	1198	808	11.9	14.6	7.98	1.383	0.0	0.056
	06/01/2012	5042	3418	10.2	15.7	8.12	0.661	0.05	0.03
	20/02/2012	ND	ND	5.9	14.7	7.91	1.895	0.058	0.03
	19/03/2012	4784	1878	10.4	15.2	7.87	9.28	0.020	0.056
Marine	14/04/2011	2313	1688	13.7	33.2	8.58	0.898	0.0	0.056
	15/06/2011	3298	1632	18.7	34.3	8.03	2.908	0.0	0.091
	13/07/2011	1909	586	21	35	8.03	5.078	0.0	0.088
	26/08/2011	ND	ND	21.8	34.2	8.23	4.48	0.0	0.095
	09/09/2011	ND	ND	20.7	35	8.07	1.946	0.0	0.062
	10/10/2011	ND	ND	18.3	35.3	7.99	2.912	0.0	0.072
	08/11/2011	ND	ND	14.5	34.8	8.03	1.312	0.0	0.261
	08/12/2011	429	337	12.6	33.6	8.02	0.803	0.0	0.072
	06/01/2012	1059	708	10.5	28.3	8.1	0.939	0.0599	0.075
	20/02/2012	914	334	6	32.8	7.95	1.419	0.071	0.082
	19/03/2012	1179	350	9.5	33.7	7.9	7.5	0.082	0.098

Supplementary Table 2. Mean number of sequences associated with each abundant OTUs retrieved in the freshwater, mesohaline and marine stations. ND: not determined.

Station	MGI Subcluster	Representative OTU	Average number of sequences		
			Apr-Sep	Oct-Nov	Dec-Mar
Freshwater station	Marine A	D32_HUJB0N002IFCRR	0	22	0
		R34_HUJB0N002JELSP	13	86	24
	Freshwater A	R16_HUJB0N002GU20W	10	62	10
		R24_HUJB0N002JCR91	6	55	8
		R8_HTRM39R02I8CB5	2	28	1
	Freshwater B	R24_HUJB0N002HQ6B7	10	26	3
		R16_HUJB0N002HSI5K	7	27	3
		R8_HTRM39R02IN9W4	0	22	2
	Sediment	R16_HUJB0N002JEY7S	15	61	9
		R10_HUJB0N002JLORV	4	30	5
		R7_HTRM39R02F36EW	4	29	3
Mesohaline station	Marine A	R22_HTRM39R02HVYCX	49	66	398
		D32_HUJB0N002F6TFA	5	19	303
		D32_HUJB0N002I2ZLT	1	6	84
		R16_HUJB0N002H3P4B	323	101	47
		R19_HUJB0N002JY578	52	18	9
		R17_HTRM39R02HMXG2	59	16	3
		R16_HUJB0N002ITPEH	175	73	24
		R16_HUJB0N002ICU9P	87	34	8
		D32_HUJB0N002IFCRR	129	131	105
		R17_HTRM39R02HNRQI	26	115	3
	Sediment	R16_HUJB0N002JEY7S	181	71	31
		R10_HUJB0N002JLORV	48	17	11
	Freshwater A	R24_HUJB0N002JCR91	48	24	18
		R16_HUJB0N002GU20W	115	43	50
		R34_HUJB0N002JELSP	199	84	78
	Freshwater B	R24_HUJB0N002HQ6B7	81	38	47
		R16_HUJB0N002HSI5K	47	19	29
Marine station	Marine A	D32_HUJB0N002F6TFA	34	ND	88
		R22_HTRM39R02HVYCX	47	ND	75
		D32_HUJB0N002I2ZLT	12	ND	18
		D32_HUJB0N002IFCRR	26	ND	12
		R16_HUJB0N002H3P4B	12	ND	7
		R25_HTRM39R02GJ5BA	10	ND	8



Supplementary Figure 1. Box plot of Shannon index associated with the three different sampling stations.



Supplementary Figure 2. Ordination diagram from CCA of active archaeal major taxonomic groups compared with environmental data.

DISCUSSION ET PERSPECTIVES

L'utilisation des techniques d'écologie moléculaire a permis au cours de ces deux dernières décennies d'accroître considérablement les connaissances relatives à la diversité du monde microbien et son rôle dans les écosystèmes. Plusieurs découvertes peuvent illustrer ces progrès technologiques telles que la mise en évidence d'*Archaea* dans les environnements aquatiques oxiques (DeLong, 1992; Fuhrman *et al.*, 1992), l'implication des *Thaumarchaeota* dans le processus d'oxydation de l'ammonium (Hatzenpichler, 2012; Schleper and Nicol, 2010) ou encore la production de méthane en zone oxygénée par des *Euryarchaeota* méthanogènes (Grossart *et al.*, 2011). A ce jour, la présence d'*Archaea* est avérée dans la zone oxygénée de la plupart des écosystèmes aquatiques tels que les milieux marins (DeLong, 1992; Galand *et al.*, 2010), les estuaires (Bernhard and Bollmann, 2010; Crump and Baross, 2000) ou encore les milieux d'eau douce (Auguet *et al.*, 2009; Pouliot *et al.*, 2009). Néanmoins, peu de données concernant la structure fine et l'activité des différentes communautés d'*Archaea* sont disponibles. En effet, les patrons de diversité des communautés microbiennes ont dû être reconsidérés, avec une richesse spécifique bien plus importante qu'imaginée en raison de la mise en évidence de la biosphère rare. D'autre part, la structure des communautés d'*Archaea* actives en lien avec les paramètres environnementaux reste encore mal connue.

L'existence de la biosphère rare archéenne est admise depuis 2009 (Galand *et al.*, 2009), cependant peu de données concernant son activité et son impact sur le fonctionnement des écosystèmes sont disponibles. Cette fraction revêt aujourd'hui un intérêt écologique majeur dans la mesure où il pourrait s'agir d'un pool de diversité génétique inexploré. D'autre part, cette fraction serait susceptible d'interagir avec d'autres communautés microbiennes, et pourrait ainsi avoir un rôle dans les successions écologiques.

La structure fine des *Archaea* pourrait aussi se traduire par la présence d'écotypes, c'est à dire de clusters phylogénétiques associés à des niches écologiques particulières. L'existence de ces écotypes chez les *Thaumarchaeota* pourrait ainsi être liée à la présence de gènes actifs spécifiques, *amoA* et/ou *ureC*, dans des conditions environnementales déterminées. De telles conditions ont pu être mises en évidence grâce à notre stratégie multi-écosystèmes présentant des gradients physico-chimiques contrastés. La dynamique de ces écotypes pourrait également être dépendante de compétitions au sein du plancton. Nous nous sommes ainsi attachés à évaluer l'activité des communautés nitrifiantes, AOA et AOB.

La discussion de cette thèse sera articulée autour de trois parties dont la première s'attachera à évaluer les apports et les biais potentiels liés à l'approche de séquençage haut-débit utilisée. Dans un second temps, deux méta-analyses conduites sur l'ensemble de ces

travaux de thèse nous permettront de discuter (i) de la diversité phylogénétique des *Archaea* actives dans les écosystèmes aquatiques, en lien avec la détection d'écotypes, et (ii) de la compétition entre AOA et AOB actives dans le processus de nitrification. Enfin, une dernière partie s'attachera à discuter de l'importance de la biosphère rare archéenne dans les écosystèmes aquatiques.

6.1 Apports et biais des études de métagénétique

La technique de pyroséquençage d'amplicons repose sur une amplification spécifique d'une région génomique cible, préalable à un séquençage massif, et au même titre que le séquençage Sanger, des biais liés à la PCR et au séquençage sont à noter. Le taux d'erreurs par base découlant du pyroséquençage est comparable à celui découlant du séquençage Sanger (Huse *et al.*, 2007) mais la différence d'échelle dans le nombre de séquences produites associé à l'absence d'un séquençage dans les deux sens, augmentent considérablement l'impact de ces erreurs sur l'étude de la diversité. Il est donc nécessaire d'établir des critères et des filtres de qualité permettant de distinguer la diversité réelle de la diversité artificielle (Kunin *et al.*, 2009; Quince *et al.*, 2009). D'autre part, l'affiliation taxonomique précise d'un grand nombre de séquences reste à ce jour un défi.

6.1.1 Spécificité des amorces

Au cours de cette thèse, l'idée était de cibler l'ensemble des *Archaea* présentes dans les écosystèmes aquatiques. Cependant, l'utilisation d'un seul couple d'amorces dites universelles pour détecter l'ensemble de la communauté archéenne n'a pas été possible. En effet, aucun couple testé ne permettait d'amplifier les communautés d'*Archaea* présentes dans les différents milieux choisis pour les différentes études et l'emploi de deux couples différents a été nécessaire.

C'est dans ce contexte que les régions hypervariables V3-V5 des gènes d'ADNr 16S ont été ciblées en milieu côtier (Chapitre 1, Article 1) et estuarien (Chapitre 2, Article 4) via l'utilisation du couple d'amorces 349F-806R (Takai and Horikoshi, 2000) tandis qu'en milieux lacustres (Chapitre 1, Article 2) le couple 519F-915R ciblant les régions hypervariables V4-V5 a été utilisé (Casamayor *et al.*, 2002; Herfort *et al.*, 2009). Ces couples d'amorces ne ciblent qu'une fraction des séquences d'*Archaea* sélectionnées dans la base

PANAM (Table 3). Cependant, comme 78 % des séquences sont amplifiées par les 2 couples d'amorce *in silico*, malgré des conditions expérimentales différentes nous avons choisi de compiler toutes les données en prenant en compte la région commune de ces séquences dans la partie 6.2 de cette discussion.

Table 3. Pourcentage de couverture du domaine des *Archaea* associé aux amplicons générés par l'utilisation des amorces 349F-806R et 519F-915R, dans la base de données *Archaea* de PANAM.

Amorce	Séquence de l'amorce (5'-3')	% de couverture du domaine <i>Archaea</i> de l'amplicon
349F	GYGCASCAGKCGMGA AW	69
806R	GGACTACVSGGGTATCTAAT	
519F	CAGCCGCCGCGGTAA	76
915R	GTGCTCCCCGCCAATTCCT	

De façon surprenante, une grande partie des séquences obtenues lors des différentes études n'a pas pu être utilisée car ces dernières étaient affiliées à des bactéries. En effet, pour cette raison, après la procédure de nettoyage, 71, 55 et 64% des séquences sont retirées des jeux de données obtenus en milieu côtier (Chapitre 1, Article 1), en milieux lacustres (Chapitre 1, Article 2) et en milieu estuarien (Chapitre 2, Article 4) respectivement. Ainsi, ces amorces sont non seulement loin d'être universelles (Table 3) mais elles manquent de spécificité. A l'avenir, un effort supplémentaire devra être fait pour dessiner des amorces pour le domaine des *Archaea* en complément de travaux réalisés précédemment (Gantner *et al.*, 2010).

Il faut donc garder à l'esprit que dans le cas du séquençage Sanger classique ou des NGS, la diversité ciblée dépend du couple d'amorce considéré. Les NGS, en raison du nombre des séquences générées, permettent en principe d'accéder à une richesse plus importante que celle évaluée par la méthode Sanger. Préalablement à toutes nos études, nous avons vérifié, grâce à l'établissement de courbes de raréfaction (établies sur les séquences affiliées aux *Archaea*), si l'effort de séquençage était suffisant pour décrire un maximum de diversité archéenne. Nous avons alors pu voir que pour la majorité des échantillons, la saturation était atteinte.

6.1.2 Nettoyage des séquences et clusterisation

Parmi les biais affectant les données de pyroséquençage, la qualité des séquences est souvent mise en exergue. En effet, certains travaux présument qu'une surévaluation de la richesse pourrait être le fruit d'erreurs de séquençage (Quince *et al.*, 2009), et/ou de l'utilisation d'un seuil de clusterisation inadapté (Huse *et al.*, 2010).

Afin de minimiser ces erreurs et d'éviter une éventuelle surestimation de la richesse (i.e. nombre d'OTUs), une procédure de nettoyage a été mise en place et décrite dans le paragraphe 3.2.4. Elle découle des travaux de thèse de Najwa Taib, qui a mis au point un pipeline d'analyse, PANAM, dédié à l'analyse des amplicons issus de NGS (Taib *et al.*, 2013). En complément de ces procédures de nettoyage, nous avons utilisé un seuil de clusterisation adapté, permettant *a priori* de ne pas surestimer la richesse. Ce seuil a été choisi en accord avec une étude antérieure (Kim *et al.*, 2011b) qui atteste qu'une similarité de 97% est optimale pour la définition des OTUs archéennes dans les régions hypervariables V3-V5.

6.1.3 Reconstruction phylogénétique

A l'heure où la deuxième génération de NGS permet de produire des séquences d'environ 400 pb, des études récentes ont montré que les méthodes phylogénétiques restituaient une taxonomie fiable (Jeraldo *et al.*, 2011; Ragan-Kelley *et al.*, 2012; Taib *et al.*, 2013). Ainsi, au cours de nos travaux nous avons pu restituer des clades précédemment définis dans la littérature par séquençage Sanger (Figure 21, Galand *et al.*, 2010).

Ceci témoigne que des séquences d'environ 400 pb peuvent être aussi informatives que des séquences complètes. La profondeur de séquençage associée à la phylogénie permet de mettre en évidence les clades C et D constitués majoritairement d'OTUs rares.

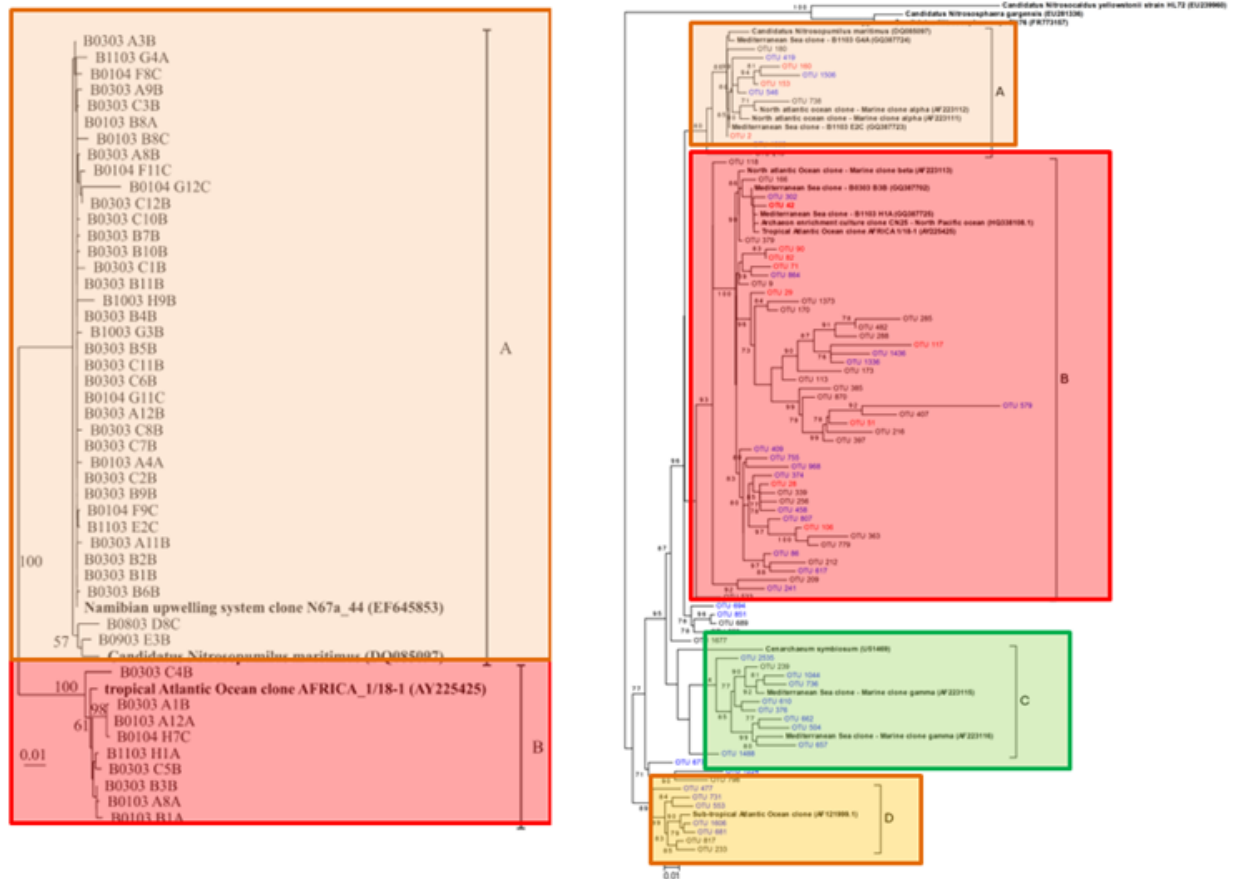


Figure 21 : Arbres phylogénétiques construits pour les *Thaumarchaeota* du Marine Group I sur la base de l'analyse des séquences des gènes d'ADNr 16S, issues d'un clonage-séquençage Sanger (à gauche, (Galand *et al.*, 2010)) et des travaux de pyroséquençage 454 sur amplicons réalisés au cours de ces travaux (à droite, Chapitre 1, Article 1). Les clades A et B sont restitués dans les deux cas, alors que seule notre étude conduite avec du pyroséquençage permet de mettre en évidence les clades C et D, constitués majoritairement d'OTUs rares.

6.2 Composition et activité de la communauté archéenne aquatique

Au cours de ce travail, nous avons pu décrire une diversité archéenne classiquement associée aux environnements aquatiques. Ainsi, nous avons détecté des groupes connus au sein des *Thaumarchaeota* (MGI) et des *Euryarchaeota* (MEG, LDS...), et dans une moindre mesure parmi les *Crenarchaeota*. Néanmoins, nos efforts ont porté sur la caractérisation de la fraction active de ces *Archaea* qui fait l'originalité de ce travail.

A cet effet, nous avons conduit une analyse (Table 4) visant à évaluer l'activité des groupes majoritaires retrouvés dans les différents écosystèmes étudiés lors de ces travaux. Une méthode utilisée pour évaluer l'activité des communautés passe par le calcul du ratio

ARN/ADN (Campbell *et al.*, 2011). Ainsi, on considère que pour un ratio supérieur à 1, le groupe ciblé est actif alors que si ce ratio est < 1 , le groupe sera considéré comme inactif.

Pour conduire cette analyse, nous avons choisi de regrouper l'écosystème côtier avec la station marine de l'estuaire nous permettant de définir les écosystèmes marins de surface oxygénés (3 m et 50 cm respectivement). Dans un second temps, nous avons considéré la station mésohaline de l'estuaire comme un milieu de salinité intermédiaire. Enfin, nous avons choisi de distinguer les écosystèmes d'eau douce oxygénés (station d'eau douce de l'estuaire, 3 m, et épilimnion/oxycline des deux lacs, 2 m et 45 m pour le lac Pavin et Bourget respectivement) et anoxiques (zone anoxique lacustres, 80 m et 140 m pour le lac Pavin et Bourget respectivement). Pour rendre cette analyse la plus pertinente possible, nous avons également introduit des données de pyroséquençage non exploitées dans les différents articles présentés dans ce manuscrit, comme celles obtenues dans l'épilimnion des écosystèmes lacustres. Ainsi, l'ensemble des données disponibles nous permettent de dégager une tendance générale sur l'activité des *Archaea* aquatiques.

Table 4. Résultats de la méta-analyse conduite sur le milieu marin (station marine de l'estuaire de la Charente et milieu côtier), milieu de salinité intermédiaire (station mésohaline de l'estuaire de la Charente) et d'eau douce oxique (station d'eau douce de l'estuaire de la Charente, épilimnions et oxyclines lacustres) et anoxique (zone anoxique lacustre). Les ratios ARN/ADN et le pourcentage d'ARN sont présentés pour les groupes taxonomiques majoritaires. La mention ND pour non déterminé traduit l'absence de données ADNr 16S pour la station mésohaline de l'estuaire de la Charente.

<i>Phylum</i>	Groupe	Milieu marin		Milieu de salinité intermédiaire		Milieu d'eau douce oxique		Milieu d'eau douce anoxique	
		Ratio	% ARN	Ratio	% ARN	Ratio	% ARN	Ratio	% ARN
<i>Euryarchaeota</i>	DHVEG-6	1,65	1,21	ND	3,59	0,31	1,96333	0,79	1,66
	MEG	1,01	0,98	ND	0,06	31,76	25,391	3,04	10,12
	LDS	0,66	0,27	ND	0,41	1,41	0,915	2,68	2,34
	<i>Methanosaeta</i>	ND	ND	ND	0,35	5,05	13,935	7,31	41,43
	RC-V	ND	0,20	ND	2,24	0,66	1,4125	0,59	0,175
	<i>Methanomicrobiales</i>	ND	ND	ND	ND	0,00	0,31	25,20	5,3405
	MGIIA	1,39	38,82	ND	2	ND	0,095	ND	ND
	MGIIIB	1,01	12,33	ND	1,53	ND	ND	ND	ND
<i>Thaumarchaeota</i>	MGI	1,33	43,05	ND	88,45	10,79	71,616	135,25	48,6
<i>Crenarchaeota</i>	C3	ND	ND	ND	0,06	0,75	0,76	9,98	14,35
	MCG	0,40	0,20	ND	ND	1,28	1,07333	1,27	2,065

L'utilisation des ARNr 16S comme indicateurs de l'activité microbienne pose cependant quelques interrogations. En effet, au sein des communautés, il existe une relation complexe entre concentration en ARN, taux de croissance et activité (Blazewicz *et al.*, 2013). Il faut donc avoir conscience que concentration en ARN et taux de croissance ne sont pas toujours corrélés de façon simple et peuvent dépendre des conditions environnementales, de même que cette relation est taxon-dépendante. En outre, il a été démontré que des cellules en dormance pouvaient avoir un contenu important de ribosomes, parfois même plus important que celui retrouvé à l'état végétatif (Blazewicz *et al.*, 2013). Néanmoins, les taux d'ARNr représentent une activité potentielle future et donnent une première information quant au statut métabolique des communautés. Son utilisation couplée à des mesures directes d'activité métabolique permettrait de palier aux biais éventuels liés au séquençage massif des ARNr 16S (Blazewicz *et al.*, 2013).

Il apparaît clairement qu'en milieu marin, peu de groupes peuvent être considérés comme actifs (i.e. avec des ratios ARN/ADN > 1), il s'agit principalement des groupes DHVEG-6, MGII.A et MGI. En milieu de salinité intermédiaire, nous n'avons pas pu établir le ratio ARN/ADN, en revanche, le pourcentage en ARN représenté pour chaque groupe suggère que les groupes les plus actifs sont MGI, et DHVEG-6. En milieu d'eau douce oxygéné, le groupe MEG a le ratio le plus élevé, suivi du MGI puis des *Methanosaeta*, alors qu'en anoxie, les groupes les plus actifs changent. En effet, on note une très nette prédominance d'activité des MGI, *Methanomicrobiales*, suivis des groupes C3 et *Methanosaeta*.

6.2.1 Ecotypes liés aux groupes MGI et MGII

La définition que nous avons considérée pour le terme écotipe lors de nos travaux est celle retenue par de nombreux auteurs dans la littérature. A savoir qu'il s'agirait de groupes de microorganismes jouant des rôles écologiques différents et faisant partie d'entités génétiques propres (groupes monophylétiques), et donc de lignées évolutives irréversiblement distinctes (Koeppel *et al.*, 2008). En d'autres termes, il s'agirait de populations moléculaires présentant une distribution unique le long de gradients écologiques (Ward, 2006). L'existence d'écotypes a déjà été prouvée, notamment pour *Prochlorococcus* (cyanobactérie marine) pour lequel différentes souches ont été séquencées. Les analyses de génomique comparative conduites sur ce microorganisme ont mis en évidence une différence phylogénétique en lien

avec l'adaptation des écotypes à la lumière, illustrant des caractéristiques physiologiques différentes (Moore *et al.*, 1998). Néanmoins, pour des microorganismes non cultivés, la différenciation des clades en différents écotypes n'est que putative et il paraît nécessaire de considérer à la fois les reconstructions phylogénétiques disponibles en lien avec les paramètres environnementaux suivis.

Sur la base de cette définition, les différentes reconstructions phylogénétiques nous ont permis de définir des clusters et des subclusters au sein même des différents groupes d'*Archaea*, à savoir MGI et MGII, retrouvés classiquement dans les écosystèmes aquatiques. Le partitionnement phylogénétique de ces différents clusters en lien avec les paramètres environnementaux suivis lors des différentes études nous a permis de supposer pour la première fois la présence d'écotypes actifs au sein des *Euryarchaeota* mais également au sein des *Thaumarchaeota*.

Il apparaît clairement qu'en milieu marin de surface, les *Euryarchaeota* du Marine Group II représentent une majorité de l'assemblage actif archéen (évalué sur la base du pourcentage de séquences ARN affiliées à ce groupe et égal à 51,15% pour ce groupe, Table 4) avec des ratios ARN/ADN compris entre 1,01 et 1,39. Plus la salinité diminue, moins le pourcentage en ARN représenté par ce groupe est important (égal à 3,53 en milieu de salinité intermédiaire) ; en milieu d'eau douce, on ne retrouve jamais ce groupe au cours de ces travaux (Table 4). Ces résultats sont en accord avec différentes études qui témoignent de l'importance quantitative de ce groupe en milieu marin (Galand *et al.*, 2010; Karner *et al.*, 2001). Ainsi, nos résultats viennent confirmer les résultats observés lors des analyses basées sur les ADNr 16S et suggèrent que la salinité serait l'un des facteurs déterminants pour l'activité (i.e. ARNr 16S) de ce groupe.

Nous avons conduit une analyse phylogénétique afin d'évaluer la proximité phylogénétique des différentes OTUs abondantes et actives de MGII en milieu marin (Figure 22). Les résultats obtenus ne témoignent pas d'une ségrégation claire des OTUs en fonction des environnements ciblés et ne permettent pas de démontrer que les populations retrouvées en milieu côtier, ou estuarien sont différentes.

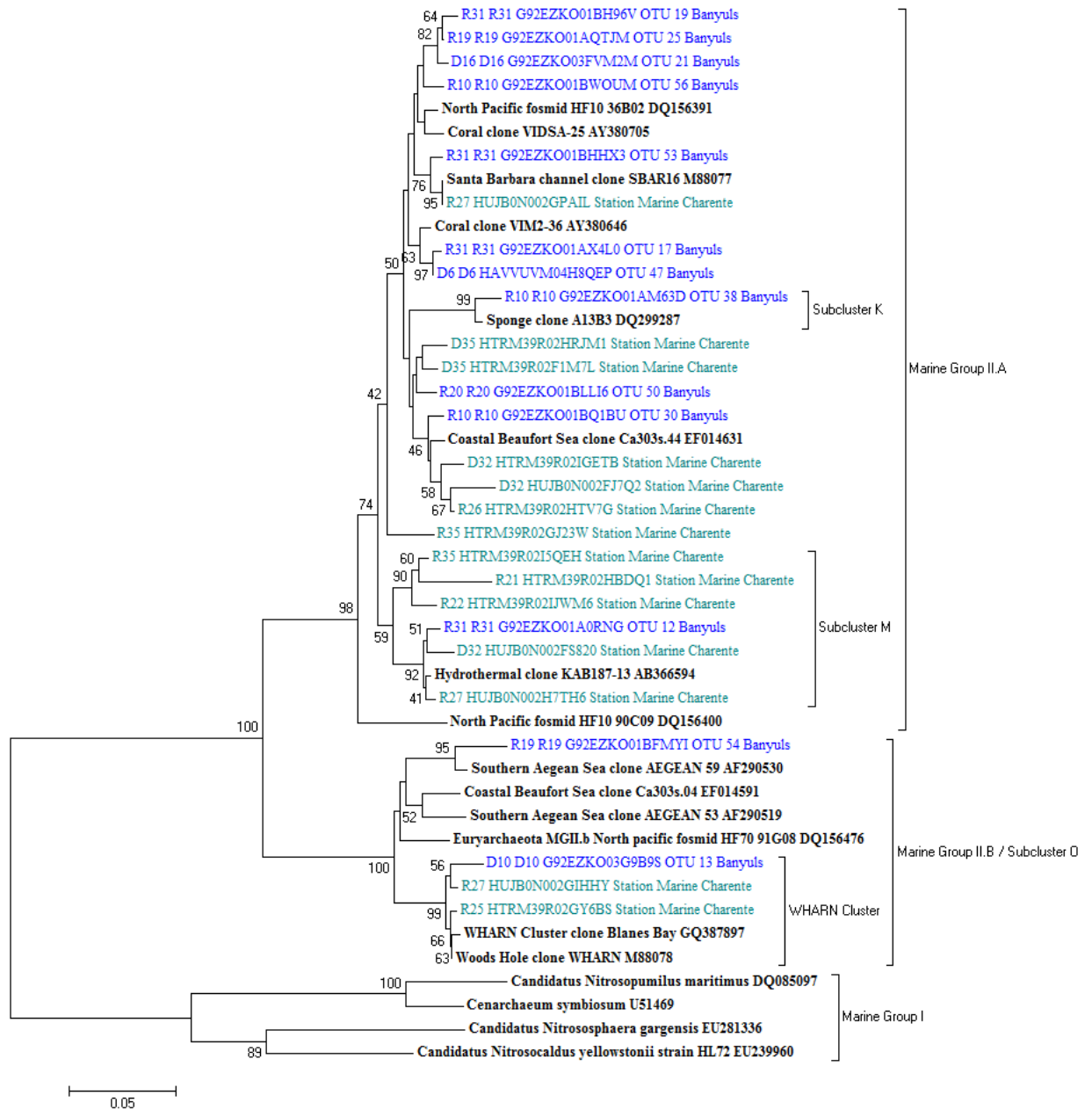


Figure 22 : Arbre phylogénétique des OTUs abondantes actives de MGII, basé sur l'analyse des séquences des gènes d'ARNr 16S, retrouvées en milieu côtier et dans la station marine de l'estuaire de la Charente. Les séquences de référence apparaissent en gras. Les séquences en vert sont issues de la station marine de l'estuaire de la Charente et les bleues du milieu côtier (Banyuls sur Mer). Les valeurs de bootstrap sont exprimées en pourcentage. Les séquences ont été alignées en utilisant Muscle (Edgar, 2004) et l'arbre a été construit en utilisant Mega5 (méthode Neighbor-Joining, 500 bootstraps) (Tamura et al., 2011).

Par le biais de l'étude conduite en milieu côtier (Chapitre 1, Article 1), nous avons pu mettre en évidence deux clusters décrits précédemment (Galand *et al.*, 2010) au sein du MGII. Par la suite nous avons mis en lien ces clusters avec différents paramètres environnementaux. L'activité du MGII.A est liée à la période estivale tandis que celle du MGII.B est liée à la

période hivernale, suggérant un effet prépondérant de la température. La récurrence saisonnière interannuelle dans l'activité de chacun de ces clusters concomitante avec des conditions environnementales reproductibles pourrait indiquer une spécialisation écologique confortant l'idée que ces clusters de MGII seraient des écotypes conditionnés par des niches écologiques différentes. La récurrence estivale du MGII.A est cohérente avec la mise en évidence de gènes codant la protéorhodopsine sur des fragments de fosmides affiliés au MGII.A dans des eaux marines (Frigaard *et al.*, 2006; Iverson *et al.*, 2012). Ceci pourrait témoigner d'une adaptation physiologique permettant une activité plus importante de ce cluster lorsque l'irradiance est plus importante. Néanmoins, en l'absence de représentants cultivés et de données physiologiques associées, ceci n'est qu'une hypothèse.

Les *Thaumarchaeota*, dont la majorité des séquences est affiliée au Marine Group I, représentent l'un des groupes archéen retrouvé classiquement dans les milieux aquatiques (Auguet *et al.*, 2011; Karner *et al.*, 2001; Pouliot *et al.*, 2009). Nous avons pu mettre en évidence des *Thaumarchaeota* actives dans tous les écosystèmes échantillonnés (représentant 43,05% des ARNr n milieu marin et jusqu'à 88,45% dans le milieu de salinité intermédiaire, Table 4). Néanmoins, l'activité de ce groupe est plus importante dans les milieux d'eau douce et de façon surprenante, en anoxie où le ratio ARN/ADN atteint 135,25 dans le lac du Bourget (Table 4).

Si l'on considère l'implication de ces microorganismes dans l'oxydation de l'ammonium en présence d'oxygène, l'activité de ce groupe en zone anoxique suscite plusieurs questions. A savoir, est ce que toutes les *Thaumarchaeota* sont réellement capables de réaliser ce processus ? Ne seraient-elles pas impliquées dans des voies métaboliques alternatives encore non décrites ? En effet, bien que le sujet soit encore débattu par la communauté scientifique, des métabolismes basés sur l'autotrophie ou encore la mixotrophie ont été évoqués (Herndl *et al.*, 2005; Ingalls *et al.*, 2006) suggérant la plasticité métabolique de ces microorganismes. Ainsi, il paraît judicieux d'envisager d'autres voies que la chimioautotrophie chez ces microorganismes.

Afin d'évaluer la distance phylogénétique entre les différentes séquences affiliées aux *Thaumarchaeota* dans les écosystèmes échantillonnés, nous avons conduit une analyse comparative de l'ensemble des OTUs de MGI abondantes retrouvées au sein des écosystèmes côtier, estuarien et des deux lacs (Figure 23). Ces résultats nous permettent de distinguer un groupe enrichi en OTUs issues des stations estuariennes mésohaline et marine au sein du cluster « Marine A », et un subcluster constitué uniquement d'OTUs issues de milieux

lacustres associé à des séquences incluses dans le subcluster « Freshwater B ». D'un autre côté, il semblerait que les séquences de MGI retrouvées en milieu estuarien mésohalin et d'eau douce soient majoritairement affiliées aux subclusters « Sédiment » et au « Freshwater A ». Cette analyse nous permet d'attester d'une différence phylogénétique des OTUs marines et d'eaux douces, qui pourrait être associée à des adaptations.

Cependant, même si certaines études ont mis en évidence une diversité phylogénétique au sein du MGI, définissant la présence d'écotypes différents (Hu *et al.*, 2011; Tully *et al.*, 2012), nous n'avons pas pu définir de clusters de MGI répondant à des paramètres environnementaux différents lors de l'étude conduite en milieu côtier (Chapitre 1, Article 1). En effet, alors que nous avons pu restituer les clusters MGI.A, B, C et D, nous n'avons pas pu mettre l'activité de ces groupes phylogénétiquement distincts en lien avec des paramètres environnementaux particuliers, ne nous permettant pas de poser l'hypothèse selon laquelle ces groupes constitueraient des écotypes. Néanmoins, nos connaissances sur l'ensemble des paramètres pouvant impacter la structuration des communautés restent limitées, et contrairement à la finesse des études moléculaires, les mesures physico-chimiques dont nous disposons peuvent paraître insuffisamment précises pour définir différentes niches écologiques.

De façon contrastée, le travail conduit sur la fraction active estuarienne (Chapitre 2, Article 4), nous a permis de suggérer la présence d'écotypes actifs de MGI. En effet, à l'échelle du subcluster cette fois, nous avons pu déterminer au sein du cluster Marine A (appelé MGI.A dans le Chapitre 1, Article 1) des dynamiques saisonnières contrastées. A cet effet, nous avons pu identifier trois subclusters dont les activités diffèrent dans le temps. En outre, la comparaison des OTUs associées à ces différents subclusters à la dynamique des transcrits *amoA*, nous permet d'affirmer ce cluster Marine A est impliqué dans l'oxydation de l'ammonium.

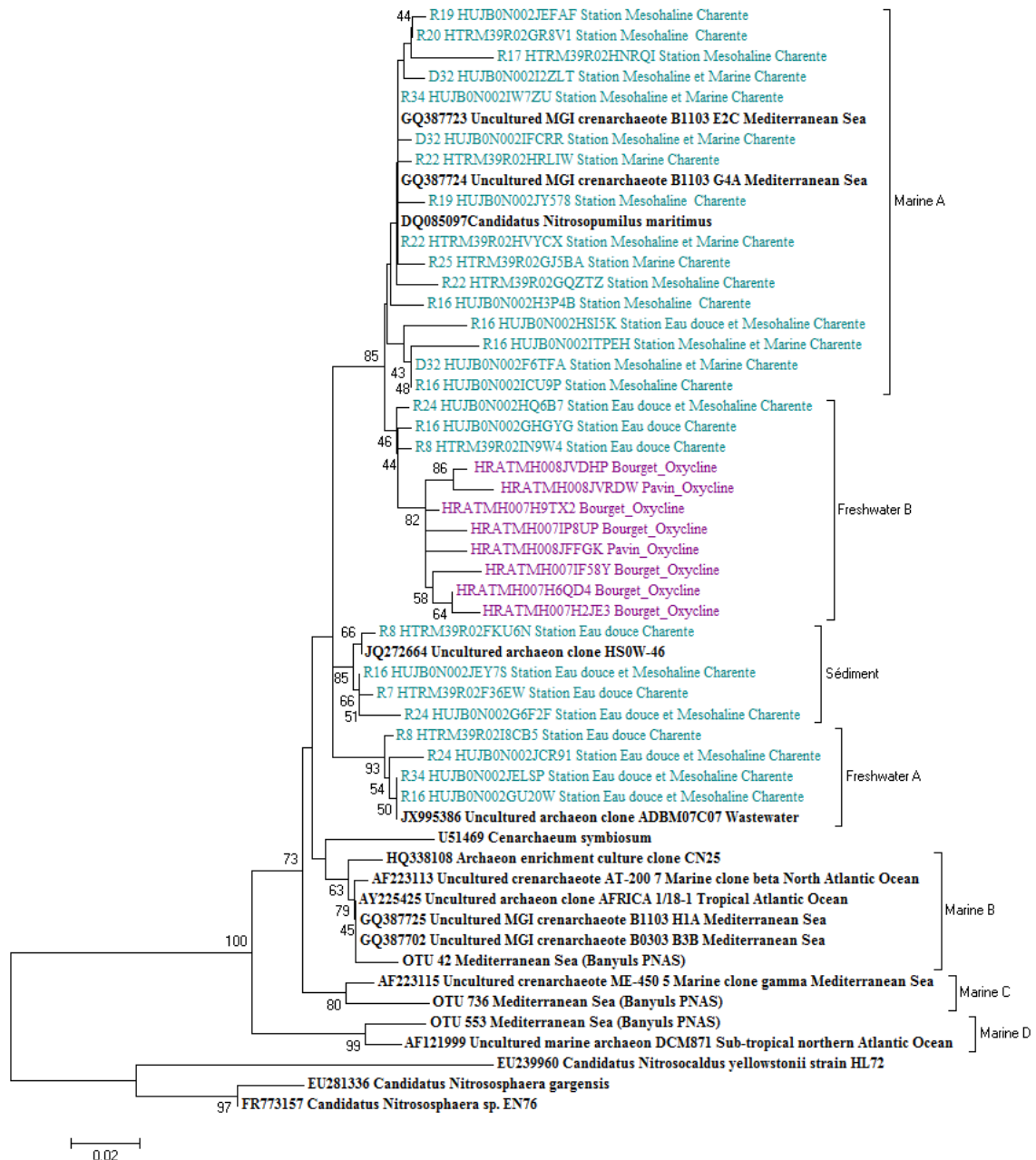


Figure 23 : Arbre phylogénétique des OTUs abondantes actives de MGI, basé sur l'analyse des séquences des gènes d'ARNr 16S, retrouvées en milieu côtier et dans la station marine de l'estuaire de la Charente. Les séquences de référence apparaissent en gras. Les séquences en vert sont issues de la station marine de l'estuaire de la Charente et les bleues du milieu côtier (Banyuls sur Mer) et en violet les séquences lacustres. Les valeurs de bootstrap sont exprimées en pourcentage. Les séquences ont été alignées en utilisant Muscle (Edgar, 2004) et l'arbre a été construit en utilisant Mega5, méthode Neighbor-Joining, 500 bootstraps (Tamura et al., 2011)).

L'originalité de cette étude était aussi de pouvoir mettre en lien le processus d'oxydation de l'ammonium avec le métabolisme de l'urée qui permettrait de fournir de l'ammonium aux AOA. Pour ce faire, nous avons analysé l'activité des gènes *amoA*

simultanément avec les gènes *ureC* et nous avons pu mettre en évidence la présence d'un subcluster au sein du cluster Freshwater A, dont l'activité annuelle connaît une prédominance estivale (Chapitre 2, Article 4). Néanmoins, sur la base de cette étude, il ne semble pas que toutes les *Thaumarchaeota* du MGI possèdent ce trait métabolique et des investigations ultérieures afin de comprendre l'importance de ce processus d'uréolyse chez les *Thaumarchaeota* sont nécessaires.

Afin d'évaluer l'implication des *Thaumarchaeota* dans l'oxydation de l'ammonium et les facteurs impliqués dans la régulation des communautés, nous avons étudié l'activité des AOA en comparaison avec celle des AOB. Pour comparer les écosystèmes, nous avons effectué une analyse statistique (forward RDA, (Borcard *et al.*, 1992)) visant à évaluer quels étaient les paramètres environnementaux conditionnant les variations du ratio AOA/AOB. Par le biais de cette analyse, nous avons pu montrer que la salinité expliquait 56,58% ($P=0,01$) et la température expliquait 32,13% ($P=0,005$) de l'activité des AOA et AOB (basée sur le nombre de transcrits quantifiés à chaque date et dans chaque écosystème). De façon contrastée, d'autres facteurs tels que le pH, la concentration en ammonium ou en phosphates n'expliquent pas significativement l'activité de ces groupes. Afin de visualiser les résultats de cette analyse statistique, nous présenterons dans la Table 5 le rapport du nombre de transcrits AOA/AOB en lien avec la salinité et la température (été vs hiver).

Table 5. Résultats de l'analyse conduite sur le nombre de transcrits *amoA* d'AOA et d'AOB, retrouvés en milieu estuarien (Chapitre 2, Article 4) et lacustre (Chapitre 1, Article 2). Le ratio du nombre de transcrits AOA/AOB est présenté pour la période hivernale et estivale associée à chaque écosystème.

	Ratio AOA/AOB	
	Hiver	Été
Eau douce Charente	0,25	0,05
Surface Pavin	0,08	0,14
Oxycline Pavin	0,01	0,03
Surface Bourget	0,72	0,03
Oxycline Bourget	1,91	0,59
Mesohaline Charente	1,13	4,94
Marine Charente	6,32	0,67

Les résultats de nos travaux tendent ainsi à montrer que l'activité des AOA est fortement impactée par la salinité des écosystèmes étudiés. En effet, les ratios AOA/AOB sont au maximum de 1,91 et 6,32 en milieu lacustre et estuarien respectivement. Ces tendances se retrouvent également le long du gradient de salinité au sein de l'estuaire, où l'on retrouve peu de transcrits dans la station d'eau douce par rapport aux deux stations plus salées (Table 5).

La salinité est l'un des facteurs majeurs qui détermine les profils de distribution des communautés microbiennes (Lozupone and Knight, 2007). Il a également été montré, sur la base d'analyses des ADN 16S, que les communautés d'AOB et leur taux de nitrification sont négativement liés à ce facteur (Bernhard and Bollmann, 2010; Rysgaard *et al.*, 1999). Ceci a également été constaté sur la base de la quantification des transcrits *amoA* d'AOB réalisée lors de ces travaux (Chapitre 2, Article 4). Néanmoins, il semblerait que l'activité des AOA, soit elle, positivement liée à la salinité et nous permet de nous interroger sur le rôle des *Thaumarchaeota*. En effet, à l'instar d'Auguet *et al.* (2011, 2012, 2013) nous mettons en évidence la présence d'AOA dans des milieux d'eau douce mais en complément de ces travaux, nous n'avons pas pu détecter une forte activité dans de tels milieux. Ces résultats suggèreraient donc que l'activité des *Thaumarchaeota* n'est pas liée uniquement à la salinité mais résulte certainement d'une combinaison de différents facteurs déterminant une niche écologique favorable pour leur activité.

L'activité hivernale de ces AOA dans la station marine est en accord avec la récurrence hivernale des *Thaumarchaeota* mise en évidence en milieu marin (Galand *et al.*, 2010; Wuchter *et al.*, 2006). Il a été suggéré que des conditions favorables aux AOA seraient liées à cette saison, facilitant la compétition pour les ressources par rapport au phytoplancton (Pitcher *et al.*, 2011). La stratification estivale créerait au contraire des conditions défavorables en terme de compétition ou de disponibilité des nutriments (Winter *et al.*, 2009), comme le montre le ratio AOA/AOB de 0,67 durant cette saison dans la station marine estuarienne (Table 5).

6.2.2 Des groupes très actifs au métabolisme peu connu

Certains groupes de méthanogènes très actifs en zone oxique ont été mis en évidence, en adéquation avec des études précédentes (Grossart *et al.*, 2011; Hirasawa *et al.*, 2008; Lliros *et al.*, 2010), alors que des groupes très actifs pour lesquels aucun trait métabolique n'est décrit ont également été détectés. C'est le cas des groupes DHVEG-6, MEG, LDS, groupe C3 ou encore MCG, qui sont tous dans un écosystème donné, actifs à un moment.

En effet, notre étude nous permet de montrer que le groupe MEG a une activité plus importante en milieu d'eau douce de surface (avec un ratio ARN/ADN de 31,76, Table 4). Ce groupe peu connu est habituellement retrouvé dans des eaux profondes (Hirayama *et al.*, 2007) mais de nombreuses séquences le caractérisant sont également issues d'écosystèmes

terrestres (Teske and Sorensen, 2008) et de sédiments marins (Takai *et al.*, 2001a). Néanmoins, il s'agit du groupe le plus actif retrouvé dans cette zone d'eau douce oxygénée, ce qui en fait un candidat de choix pour des études ultérieures.

Dans la zone d'eau douce anoxique, les *Crenarchaeota* du groupe C3 sont également relativement actives (ratio ARN/ADN de 9,98). Ce groupe d'*Archaea* encore dépourvu de représentant cultivé, est souvent associé à des environnements terrestres et marins (DeLong and Pace, 2001), mais aussi à des environnements sédimentaires marins profonds (Wang *et al.*, 2010) et des lacs salés (Comeau *et al.*, 2012; Schneider *et al.*, 2013). D'autres groupes tels que le LDS sont retrouvés de manière minoritaire dans l'ensemble des milieux échantillonnés (Table 4). On remarque un ratio ARN/ADN de 2,68 pour ce groupe en zone anoxique lacustre, signifiant son activité. La bibliographie fait état d'une diversification importante de ce groupe, suggérant un rôle fonctionnel clé dans les milieux d'eau douce (Auguet *et al.*, 2009). Néanmoins, l'absence de représentant cultivé ou de génomes séquencés, ne permet pas d'apporter des réponses plus précises quant à son rôle dans ces milieux (Barberan *et al.*, 2011).

Ces résultats témoignent de l'importance de prendre en considération les groupes actifs les moins connus, comme le groupe C3 ou MEG, dont on ne peut plus désormais négliger la contribution au fonctionnement des écosystèmes. Ainsi, l'exploration de ces communautés pourrait constituer la suite logique de ces travaux. Elle pourrait passer par la détermination d'amorces spécifiques de ces différents *taxa* ciblant les gènes d'ADNr 16S, qui pourraient servir à développer une approche de capture de gènes (Denonfoux *et al.*, 2013). Dans le cas envisagé, cette technique, consisterait à cibler les gènes d'ADNr 16S et de capturer de grands fragments de sorte que si les gènes d'ADNr 16S sont co-localisés avec des gènes fonctionnels, on peut imaginer capturer un gène fonctionnel apportant des informations sur le potentiel métabolique de ces microorganismes, et permettant d'adapter les conditions pour leur mise en culture et leur isolement.

6.3 Importance de la biosphère rare archéenne dans les écosystèmes aquatiques

La détermination de la structure des communautés (abondance, richesse, composition) d'un écosystème est un enjeu central en écologie et donc en écologie microbienne. Dans les écosystèmes aquatiques, jusqu'à récemment plus de 90% des librairies estimaient à moins de

200 le nombre de phylotypes bactériens (OTUs). Les travaux pionniers de Sogin *et al.* (2006), basés sur les nouvelles techniques de séquençage ont permis de détecter entre 1184 et 3290 OTUs en milieu marin, suggérant ainsi que la richesse spécifique détectée jusqu'alors dans ces écosystèmes était certainement sous-estimée. En outre, la plus grande partie de la biodiversité est représentée par un grand nombre d'OTUs faiblement représentés, constituant la biosphère rare.

Une des nouveautés apportées par nos travaux vient de la description de nouveaux clades au sein des *Thaumarchaeota* du Marine Group I (Chapitre 1, Article 1). En effet, les phylogénies générées nous ont permis de confirmer l'existence du clade MGI.C, décrit brièvement au cours d'une étude conduite en milieu côtier (Massana *et al.*, 2000). Enfin, nous avons pu mettre en évidence la présence d'un nouveau clade, le MGI.D (Figure 21), montrant bien l'apport des NGS au décryptage de la diversité archéenne. Ces deux derniers sont par ailleurs caractérisés par une grande majorité d'OTUs qui sont toujours rares dans ce milieu côtier.

L'un des enjeux de ce travail était de montrer quelle pouvait être la structuration de cette biosphère rare archéenne et de répondre aux questions concernant son activité. Nous avons pu mettre en évidence un patron de diversité de la biosphère rare archéenne plus complexe que celle décrite dans la littérature jusqu'alors (Chapitre 1, Article 1). En effet, trois fractions distinctes de microorganismes rares semblent composer l'assemblage archéen dans les eaux de surface Méditerranéennes :

- Deux fractions dont la composition taxonomique est celle retrouvée pour les microorganismes abondants :
 - dont l'une est active et probablement nécessaire au fonctionnement de l'écosystème ;
 - dont l'autre est en dormance attendant des conditions environnementales optimales (température, oxygène, nutriments...) pour leur croissance et leur développement ;
- Une fraction dont la composition taxonomique est très différente de celle couramment retrouvée dans le milieu pélagique échantillonné. En outre, cette fraction considérée comme étrangère au milieu est inactive et pourrait provenir d'écosystèmes sédimentaires et profonds, arrivée dans la zone pélagique par des phénomènes de dispersion (tempêtes, courants marins...).

Les détracteurs du concept de biosphère rare considéreront certainement que la mise en évidence de cette dernière fraction de microorganismes rares et inactifs pourrait être le fruit de biais méthodologiques décrits en partie dans le paragraphe 6.1. Au cours de notre étude, nous avons choisi de conserver les singletons pour l'analyse de la biosphère rare, et afin de bien cerner la validité de nos résultats, nous avons conduit une analyse visant à évaluer quelle était la pertinence et la réalité de ces OTUs représentés par une séquence unique. Il en ressort que sur les 81 OTUs singleton du jeu de données normalisé nous ayant servi à déterminer l'activité de la biosphère rare, 75% sont retrouvées dans les deux jeux de données ADN et ARN, donc représentées par plus d'une séquence, et au même titre retrouvées à plus d'une séquence dans le jeu de données brut, il ne s'agirait donc pas d'un biais provenant du pyroséquençage, mais bien d'une diversité d'OTUs faiblement représentées. Une des façons de prouver que ces microorganismes ne sont pas des biais liés au pyroséquençage serait de développer des amorces spécifiques de ces groupes permettant de les cibler. Il s'agirait pour un groupe actif donné, d'évaluer si des motifs consensus sont retrouvés dans les données de pyroséquençage permettant la détermination d'une amorce (Lynch *et al.*, 2012). Ensuite, une amorce généraliste, serait utilisée en complément pour amplifier un fragment d'une longueur suffisante pour restituer une phylogénie. Ainsi, ces amorces spécifiques permettraient d'évaluer dans un premier temps la diversité de ces microorganismes rares, puis envisager le développement d'amorces générant des fragments plus courts permettant une quantification par PCR pour évaluer la dynamique de ces groupes.

Il n'en reste pas moins que même si 25% des séquences identifiées comme étant potentiellement des erreurs de séquençage résiduelles, 75% des séquences ne seraient pas biaisées. Ainsi, ces microorganismes rares et inactifs représenteraient une « Seed Bank » non locale, arrivée dans le milieu pélagique de surface par des phénomènes d'immigration et représenterait donc un pool de colonisateurs potentiels venus d'un autre écosystème, plus profond voir sédimentaire.

Ainsi, sur la base des travaux de Pedros-Alio (2006, 2012), nous pouvons proposer un schéma conceptuel focalisé sur l'activité des *Archaea* rares. A cet effet, nous avons établi une courbe rang-abondance basée sur le nombre de séquences ARN retrouvées pour les différents rangs d'OTUs. On note la présence d'OTUs abondantes actives dans le milieu, mais également des trois différentes fractions de microorganismes rares : ceux suivant l'hypothèse de seed bank, les *taxa* rares dont certains sont actifs et d'autres qui sont inactifs. Les *taxa* rares actifs présentent une forte identité avec les bases de données et leur affiliation taxonomique est proche de celle des *taxa* abondants. Ils interviendraient donc dans le

fonctionnement des écosystèmes mais leur maintien à l'état rare suscite encore de multiples questions, comme par exemple de savoir s'il s'agit du résultat d'une sensibilité plus importante à la lyse virale ou à la prédation. Les *taxa* rares inactifs seraient une seed bank non locale, c'est-à-dire des *Archaea* rares venues d'un autre milieu que l'écosystème de surface échantillonné, par des phénomènes d'immigration. Cependant, leur provenance, leur statut et leur rôle dans l'écosystème ne peuvent à l'heure actuelle demeurer uniquement au stade d'hypothèses. On peut donc se poser la question de savoir si le schéma dégagé pour les *Archaea* est extrapolable aux microorganismes affiliés aux bactéries ou aux eucaryotes ?

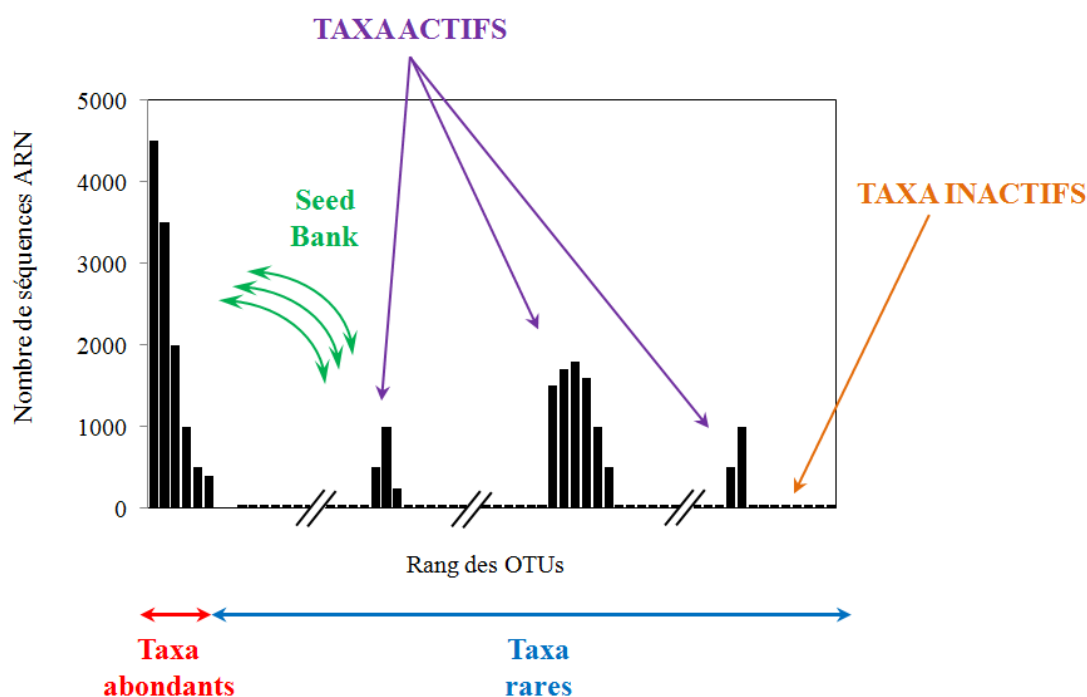


Figure 25. Schéma de synthèse basé sur l'activité des *Archaea* rares (16S rRNA) retrouvées en milieu côtier de surface. Le rang des OTUs retenu pour établir la courbe est celui retrouvé en ADN, nous permettant d'attester des variations du nombre de séquences ARN pour des *taxa* retrouvés à l'état rare dans la fraction ADN. On retrouve des microorganismes rares suivant la théorie de la seed bank, c'est-à-dire rares mais susceptibles de devenir abondants et actifs sous l'effet de conditions environnementales changeantes. On note également la présence de *taxa* rares dont certains sont actifs et d'autres inactifs.

En conclusion, ces travaux ont permis la mise en évidence d'*Archaea* rares mais néanmoins actives dans les milieux aquatiques, de même que la présence de groupes archéens actifs à la physiologie inconnue. Cet ensemble s'inscrit dans une démarche du décryptage du rôle des *Archaea* dans le fonctionnement des écosystèmes et la caractérisation ultérieure de certains de ces groupes d'intérêt constitue un véritable challenge en écologie microbienne.

REFERENCES BIBLIOGRAPHIQUES

- Abell GC, Revill AT, Smith C, Bissett AP, Volkman JK, Robert SS (2010). Archaeal ammonia oxidizers and nirS-type denitrifiers dominate sediment nitrifying and denitrifying populations in a subtropical macrotidal estuary. *ISME J* **4**: 286-300.
- Agogué H, Brink M, Dinasquet J, Herndl GJ (2008). Major gradients in putatively nitrifying and non-nitrifying *Archaea* in the deep North Atlantic. *Nature* **456**: 788-91.
- Alonso-Saez L, Balague V, Sa EL, Sanchez O, Gonzalez JM, Pinhassi J *et al* (2007). Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiol Ecol* **60**: 98-112.
- Alonso-Saez L, Waller AS, Mende DR, Bakker K, Farnelid H, Yager PL *et al* (2012). Role for urea in nitrification by polar marine *Archaea*. *Proc Natl Acad Sci USA* **109**: 17989-94.
- Arp DJ, Chain PS, Klotz MG (2007). The impact of genome analyses on our understanding of ammonia-oxidizing *bacteria*. *Annu Rev Microbiol* **61**: 503-28.
- Arrigo KR (2005). Marine microorganisms and global nutrient cycles. *Nature* **437**: 349-55.
- Auguet JC, Barberan A, Casamayor EO (2009). Global ecological patterns in uncultured *Archaea*. *ISME J* **4**: 182-90.
- Auguet JC, Casamayor EO (2008). A hotspot for cold *crenarchaeota* in the neuston of high mountain lakes. *Environ Microbiol* **10**: 1080-6.
- Auguet JC, Casamayor EO (2013). Partitioning of *Thaumarchaeota* populations along environmental gradients in high mountain lakes. *FEMS Microbiol Ecol* **84**: 154-64.
- Auguet JC, Nomokonova N, Camarero L, Casamayor EO (2011). Seasonal changes of freshwater ammonia-oxidizing archaeal assemblages and nitrogen species in oligotrophic alpine lakes. *Appl Environ Microbiol* **77**: 1937-45.
- Auguet JC, Triado-Margarit X, Nomokonova N, Camarero L, Casamayor EO (2012). Vertical segregation and phylogenetic characterization of ammonia-oxidizing *Archaea* in a deep oligotrophic lake. *ISME J* **6**: 1786-97.
- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad TF (1983). The ecological role of water column microbes in the sea. *Marine Ecology Progress Series* **10**: 257-263.
- Baas-Becking LGM (1934). Geobiologie of Inleiding Tot de Milieukunde. *serie 18/19 Van Stockum's Gravenhange*.
- Baker BJ, Lesniewski RA, Dick GJ (2012). Genome-enabled transcriptomics reveals archaeal populations that drive nitrification in a deep-sea hydrothermal plume. *ISME J* **6**: 2269-79.
- Barberan A, Fernandez-Guerra A, Auguet JC, Galand PE, Casamayor EO (2011). Phylogenetic ecology of widespread uncultured clades of the Kingdom *Euryarchaeota*. *Mol Ecol* **20**: 1988-96.

- Barns SM, Delwiche CF, Palmer JD, Pace NR (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* **93**: 9188-93.
- Bartossek R, Spang A, Weidler G, Lanzen A, Schleper C (2012). Metagenomic analysis of ammonia-oxidizing archaea affiliated with the soil group. *Front Microbiol* **3**: 208.
- Beman JM, Popp BN, Francis CA (2008). Molecular and biogeochemical evidence for ammonia oxidation by marine *Crenarchaeota* in the Gulf of California. *ISME J* **2**: 429-41.
- Beman JM, Sachdeva R, Fuhrman JA (2010). Population ecology of nitrifying archaea and bacteria in the Southern California Bight. *Environ Microbiol* **12**: 1282-92.
- Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hugler M *et al* (2010). Autotrophic carbon fixation in *Archaea*. *Nat Rev Microbiol* **8**: 447-60.
- Bernhard AE, Bollmann A (2010). Estuarine nitrifiers: New players, patterns and processes. *Estuarine, Coastal and Shelf Science* **88**: 1-11.
- Bernhard AE, Tucker J, Giblin AE, Stahl DA (2007). Functionally distinct communities of ammonia-oxidizing bacteria along an estuarine salinity gradient. *Environ Microbiol* **9**: 1439-47.
- Biddle JF, Lipp JS, Lever MA, Lloyd KG, Sorensen KB, Anderson R *et al* (2006). Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl Acad Sci USA* **103**: 3846-51.
- Blainey PC, Mosier AC, Potanina A, Francis CA, Quake SR (2011). Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS One* **6**: e16626.
- Blazewicz SJ, Barnard RL, Daly RA, Firestone MK (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J*.
- Bloch E, Rachel R, Burggraf S, Hafenbradl D, Jannasch HW, Stetter KO (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of *Archaea*, extending the upper temperature limit for life to 113 degrees C. *Extremophiles* **1**: 14-21.
- Borcard D, Legendre P, Drapeau P (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045-1055.
- Borrel G, Lehours AC, Crouzet O, Jezequel D, Rockne K, Kulczak A *et al* (2012). Stratification of *Archaea* in the deep sediments of a freshwater meromictic lake: vertical shift from methanogenic to uncultured archaeal lineages. *PLoS One* **7**: e43346.
- Bosshard PP, Santini Y, Gruter D, Stettler R, Bachofen R (2000). Bacterial diversity and community composition in the chemocline of the meromictic alpine Lake Cadagno as revealed by 16S rDNA analysis. *FEMS Microbiol Ecol* **31**: 173-182.

- Bothe H, Jost G, Schlöter M, Ward BB, Witzel K (2000). Molecular analysis of ammonia oxidation and denitrification in natural environments. *FEMS Microbiol Rev* **24**: 673-90.
- Boucher D, Richardot M, Thenot A, Debroas D (2006). Incorporation of 3H-thymidine by different prokaryotic groups in relation to temperature and nutrients in a lacustrine ecosystem. *Microb Ecol* **52**: 399-407.
- Bouskill NJ, Eveillard D, Chien D, Jayakumar A, Ward BB (2012). Environmental factors determining ammonia-oxidizing organism distribution and diversity in marine environments. *Environ Microbiol* **14**: 714-29.
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008). Mesophilic *Crenarchaeota*: proposal for a third archaeal phylum, the *Thaumarchaeota*. *Nat Rev Microbiol* **6**: 245-52.
- Brochier-Armanet C, Forterre P, Gribaldo S (2011). Phylogeny and evolution of the *Archaea*: one hundred genomes later. *Curr Opin Microbiol* **14**: 274-81.
- Brown JR, Doolittle WF (1997). *Archaea* and the prokaryote-to-eukaryote transition. *Microbiol Mol Biol Rev* **61**: 456-502.
- Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011). Activity of abundant and rare *bacteria* in a coastal ocean. *Proc Natl Acad Sci USA* **108**: 12776-81.
- Campbell BJ, Yu L, Straza TRA, Kirchman DL (2009). Temporal changes in bacterial rRNA and rRNA genes in Delaware (USA) coastal waters. *Aqua Microbial Ecol* **57**: 123-135.
- Caporaso JG, Paszkiewicz K, Field D, Knight R, Gilbert JA (2012). The Western English Channel contains a persistent microbial seed bank. *ISME J* **6**: 1089-93.
- Carbonero F, Oakley BB, Hawkins RJ, Purdy KJ (2012). Genotypic distribution of a specialist model microorganism, *Methanosaeta*, along an estuarine gradient: does metabolic restriction limit niche differentiation potential? *Microb Ecol* **63**: 856-64.
- Chapron E, Alberic P, Jezequel D, Versteeg W, Bourdier J, Sitbon J (2010). Multidisciplinary characterisation of sedimentary processes in a recent maar lake (Lake Pavin, French Massif Central) and implication for natural hazards. *Natural Hazards and Earth System Sciences* **10**: 1815-1827.
- Church MJ, DeLong EF, Ducklow HW, Karner MB, Preston CM, Karl DM (2003). Abundance and distribution of planktonic *Archaea* and *Bacteria* in the waters west of the Antarctic Peninsula. *American Society of Limnology and Oceanography* **48**: 1893-1902.
- Church MJ, Wai B, Karl DM, DeLong EF (2009). Abundances of crenarchaeal *amoA* genes and transcripts in the Pacific Ocean. *Environ Microbiol*.
- Coleman ML, Chisholm SW (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634-9.

- Colombet J, Charpin M, Robin A, Portelli C, Amblard C, Cauchie HM *et al* (2009). Seasonal depth-related gradients in virioplankton: standing stock and relationships with microbial communities in Lake Pavin (France). *Microb Ecol* **58**: 728-36.
- Comeau AM, Harding T, Galand PE, Vincent WF, Lovejoy C (2012). Vertical distribution of microbial communities in a perennially stratified Arctic lake with saline, anoxic bottom waters. *Sci Rep* **2**: 604.
- Crump BC, Baross JA (2000). Archaeaplankton in the Columbia River, its estuary and the adjacent coastal ocean, USA. *FEMS Microbiol Ecol* **31**: 231-239.
- De Boer W, Gunnewiek PJ, Veenhuis M, Bock E, Laanbroek HJ (1991). Nitrification at Low pH by Aggregated Chemolithotrophic Bacteria. *Appl Environ Microbiol* **57**: 3600-4.
- de la Torre JR, Walker CB, Ingalls AE, Konneke M, Stahl DA (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* **10**: 810-8.
- DeLong EF (1992). *Archaea* in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685-9.
- DeLong EF (1998). Everything in moderation: *archaea* as 'non-extremophiles'. *Curr Opin Genet Dev* **8**: 649-54.
- DeLong EF, Karl DM (2005). Genomic perspectives in microbial oceanography. *Nature* **437**: 336-42.
- DeLong EF, Pace NR (2001). Environmental diversity of *bacteria* and *archaea*. *Syst Biol* **50**: 470-8.
- Denonfoux J, Parisot N, Dugat-Bony E, Biderre-Petit C, Boucher D, Morgavi DP *et al* (2013). Gene capture coupled to high-throughput sequencing as a strategy for targeted metagenome exploration. *DNA Res* **20**: 185-96.
- Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-7.
- Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-1.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-200.
- Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L *et al* (2008). A korarchaeal genome reveals insights into the evolution of the *Archaea*. *Proc Natl Acad Sci USA* **105**: 8102-7.
- Erguder TH, Boon N, Wittebolle L, Marzorati M, Verstraete W (2009). Environmental factors shaping the ecological niches of ammonia-oxidizing *archaea*. *FEMS Microbiol Rev* **33**: 855-69.

- Fegatella F, Lim J, Kjelleberg S, Cavicchioli R (1998). Implications of rRNA operon copy number and ribosome content in the marine oligotrophic ultramicrobacterium *Sphingomonas* sp. strain RB2256. *Appl Environ Microbiol* **64**: 4433-8.
- Finlay BJ (2002). Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061-3.
- Forterre P, Brochier C, Philippe H (2002). Evolution of the *Archaea*. *Theor Popul Biol* **61**: 409-22.
- Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB (2005). Ubiquity and diversity of ammonia-oxidizing *archaea* in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* **102**: 14683-8.
- Frigaard NU, Martinez A, Mincer TJ, DeLong EF (2006). Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and *Archaea*. *Nature* **439**: 847-50.
- Fuhrman JA, McCallum K, Davis AA (1992). Novel major archaebacterial group from marine plankton. *Nature* **356**: 148-9.
- Galand PE, Bourrain M, De Maistre E, Catala P, Desdevises Y, Elifantz H *et al* (2012). Phylogenetic and functional diversity of *Bacteria* and *Archaea* in a unique stratified lagoon, the Clipperton atoll (N Pacific). *FEMS Microbiol Ecol*.
- Galand PE, Casamayor EO, Kirchman DL, Lovejoy C (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci USA* **106**: 22427-32.
- Galand PE, Gutierrez-Provecho C, Massana R, Gasol J, Casamayor EO (2010). Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea. *Limnol Oceanogr* **55**: 2117-2125.
- Galand PE, Lovejoy C, Pouliot J (2008). Microbial community diversity and heterotrophic production in a coastal Arctic ecosystem: A stamukhi lake and its source waters. *Limnol Oceanogr* **53**: 813-823.
- Galand PE, Lovejoy C, Vincent WF (2006). Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquatic Microbial Ecology* **44**: 115-126.
- Gantner S, Andersson AF, Alonso-Saez L, Bertilsson S (2010). Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. *J Microbiol Methods* **84**: 12-8.
- Garcia-Martinez J, Rodriguez-Valera F (2000). Microdiversity of uncultured marine prokaryotes: the SAR11 cluster and the marine *Archaea* of Group I. *Environ Microbiol* **9**: 935-948.
- Glass JB, Orphan VJ (2012). Trace metal requirements for microbial enzymes involved in the production and consumption of methane and nitrous oxide. *Front Microbiol* **3**: 61.

- Glissman K, Chin KJ, Casper P, Conrad R (2004). Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature. *Microb Ecol* **48**: 389-99.
- Glockner FO, Fuchs BM, Amann R (1999). Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization. *Appl Environ Microbiol* **65**: 3721-6.
- Godfrey LV, Glass JB (2011). The geochemical record of the ancient nitrogen cycle, nitrogen isotopes, and metal cofactors. *Methods Enzymol* **486**: 483-506.
- Grossart HP, Frindte K, Dziallas C, Eckert W, Tang KW (2011). Microbial methane production in oxygenated water column of an oligotrophic lake. *Proc Natl Acad Sci U S A* **108**: 19657-61.
- Großkopf R, Stubner S, Liesack W (1998). Novel Euryarchaeotal Lineages Detected on Rice Roots and in the Anoxic Bulk Soil of Flooded Rice Microcosms. *Appl Environ Microbiol* **64**: 4983-4989.
- Gubry-Rangin C, Hai B, Quince C, Engel M, Thomson BC, James P *et al* (2011). Niche specialization of terrestrial archaeal ammonia oxidizers. *Proc Natl Acad Sci USA* **108**: 21206-11.
- Gubry-Rangin C, Nicol GW, Prosser JI (2010). *Archaea* rather than *bacteria* control nitrification in two agricultural acidic soils. *FEMS Microbiol Ecol* **74**: 566-74.
- Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J *et al* (2006a). Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* **103**: 18296-301.
- Hallam SJ, Mincer TJ, Schleper C, Preston CM, Roberts K, Richardson PM *et al* (2006b). Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine *Crenarchaeota*. *PLoS Biol* **4**: e95.
- Hansman RL, Griffin S, Watson JT, Druffel ER, Ingalls AE, Pearson A *et al* (2009). The radiocarbon signature of microorganisms in the mesopelagic ocean. *Proc Natl Acad Sci USA* **106**: 6513-8.
- Hatzenpichler R (2012). Diversity, physiology and niche differentiation of ammonia-oxidizing archaea. *Appl Environ Microbiol*.
- Hatzenpichler R, Lebedeva EV, Spieck E, Stoecker K, Richter A, Daims H *et al* (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci USA* **105**: 2134-9.
- Herfort L, Kim JH, Coolen MJL, Abbas B, Schouten S, Herndl GJ *et al* (2009). Diversity of *Archaea* and detection of crenarchaeotal *amoA* genes in the river Rhine and Têt. *Aquatic Microbial Ecology* **55**: 189-201.

- Herlemann DP, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ, Andersson AF (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* **5**: 1571-9.
- Herndl GJ, Reinthaler T, Teira E, van Aken H, Veth C, Pernthaler A *et al* (2005). Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303-9.
- Herrmann M, Saunders AM, Schramm A (2008). *Archaea* dominate the ammonia-oxidizing community in the rhizosphere of the freshwater macrophyte *Littorella uniflora*. *Appl Environ Microbiol* **74**: 3279-83.
- Hirasawa JS, Sarti A, Del Aguila NK, Varesche MB (2008). Application of molecular techniques to evaluate the methanogenic *archaea* and anaerobic *bacteria* in the presence of oxygen with different COD:sulfate ratios in a UASB reactor. *Anaerobe* **14**: 209-18.
- Hirayama H, Sunamura M, Takai K, Nunoura T, Noguchi T, Oida H *et al* (2007). Culture-dependent and -independent characterization of microbial communities associated with a shallow submarine hydrothermal system occurring within a coral reef off Taketomi Island, Japan. *Appl Environ Microbiol* **73**: 7642-56.
- Hu A, Jiao N, Zhang R, Yang Z (2011). Niche partitioning of marine group I *Crenarchaeota* in the euphotic and upper mesopelagic zones of the East China Sea. *Appl Environ Microbiol* **77**: 7469-78.
- Hu A, Yao T, Jiao N, Liu Y, Yang Z, Liu X (2010). Community structures of ammonia-oxidizing *archaea* and *bacteria* in high-altitude lakes on the Tibetan Plateau. *Freshwater Biology*.
- Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO (2002). A new phylum of *Archaea* represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63-7.
- Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA *et al* (2007). Microbial population structures in the deep marine biosphere. *Science* **318**: 97-100.
- Hugoni M, Etien S, Bourges A, Lepere C, Domaizon I, Mallet C *et al* (2013a). Dynamics of ammonia-oxidizing *Archaea* and *Bacteria* in contrasted freshwater ecosystems. *Res Microbiol* **164**: 360-70.
- Hugoni M, Taib N, Debroas D, Domaizon I, Jouan Dufournel I, Bronner G *et al* (2013b). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci USA* **110**: 6004-9.
- Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8**: R143.
- Huse SM, Welch DM, Morrison HG, Sogin ML (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889-98.

- Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever M, Lauer A *et al* (2006). Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proc Natl Acad Sci USA* **103**: 2815-20.
- Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K *et al* (2003). Microbial communities associated with geological horizons in coastal subseafloor sediments from the sea of Okhotsk. *Appl Environ Microbiol* **69**: 7224-35.
- Ingalls AE, Shah SR, Hansman RL, Aluwihare LI, Santos GM, Druffel ER *et al* (2006). Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc Natl Acad Sci USA* **103**: 6442-7.
- Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV (2012). Untangling genomes from metagenomes: revealing an uncultured class of marine *Euryarchaeota*. *Science* **335**: 587-90.
- Jeraldo P, Chia N, Goldenfeld N (2011). On the suitability of short reads of 16S rRNA for phylogeny-based analyses in environmental surveys. *Environ Microbiol* **13**: 3000-9.
- Jones SE, Lennon JT (2010). Dormancy contributes to the maintenance of microbial diversity. *Proc Natl Acad Sci USA* **107**: 5881-6.
- Jung MY, Park SJ, Min D, Kim JS, Rijpstra WI, Sinninghe Damste JS *et al* (2011). Enrichment and characterization of an autotrophic ammonia-oxidizing archaeon of mesophilic crenarchaeal group I.1a from an agricultural soil. *Appl Environ Microbiol* **77**: 8635-47.
- Junier P, Molina V, Dorador C, Hadas O, Kim OS, Junier T *et al* (2010). Phylogenetic and functional marker genes to study ammonia-oxidizing microorganisms (AOM) in the environment. *Appl Microbiol Biotechnol* **85**: 425-40.
- Jurgens G (2000). Molecular phylogeny of *Archaea* in boreal forest soil, freshwater and temperate estuarine sediment. *Thèse*.
- Karner MB, DeLong EF, Karl DM (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-10.
- Kemp PF, Aller JY (2004). Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol Ecol* **47**: 161-77.
- Keough BP, Schmidt TM, Hicks RE (2003). Archaeal nucleic acids in picoplankton from great lakes on three continents. *Microb Ecol* **46**: 238-48.
- Kim BK, Jung MY, Yu DS, Park SJ, Oh TK, Rhee SK *et al* (2011a). Genome sequence of an ammonia-oxidizing soil archaeon, "Candidatus *Nitrosoarchaeum koreensis*" MY1. *J Bacteriol* **193**: 5539-40.
- Kim M, Morrison M, Yu Z (2011b). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods* **84**: 81-7.

- Kirchman DL, Dittel AI, Malmstrom RR, Cottrell MT (2005). Biogeography of major bacterial groups in the Delaware Estuary. *Limnol Oceanogr* **50**: 1697-1706.
- Klotz MG (2011). Methods in Enzymology. Research on Nitrification and Related Processes, Part A. Preface. *Methods Enzymol* **486**: xix-xx.
- Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM *et al* (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci USA* **105**: 2504-9.
- Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543-6.
- Konstantinidis KT, Braff J, Karl DM, DeLong EF (2009). Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific subtropical gyre. *Appl Environ Microbiol* **75**: 5345-55.
- Kowalchuk GA, Stephen JR (2001). Ammonia-oxidizing bacteria: a model for molecular microbial ecology. *Annu Rev Microbiol* **55**: 485-529.
- Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2009). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12**: 118-23.
- Kuypers MM, Slikers AO, Lavik G, Schmid M, Jorgensen BB, Kuenen JG *et al* (2003). Anaerobic ammonium oxidation by anammox bacteria in the Black Sea. *Nature* **422**: 608-11.
- La Cono V, La Spada G, Arcadi E, Placenti F, Smedile F, Ruggeri G *et al* (2013). Partaking of *Archaea* to biogeochemical cycling in oxygen-deficient zones of meromictic saline Lake Faro (Messina, Italy). *Environ Microbiol* **15**: 1717-33.
- Lami R, Ghiglione JF, Desdevises JF, West NJ, Lebaron P (2009). Annual patterns of presence and activity of marine bacteria monitored by 16S rDNA-16SrRNA fingerprints in the coastal NW Mediterranean Sea. *Aquat Microb Ecol* **54**: 199-210.
- Lehours AC, Bardot C, Thenot A, Debroas D, Fonty G (2005). Anaerobic microbial communities in Lake Pavin, a unique meromictic lake in France. *Appl Environ Microbiol* **71**: 7389-400.
- Lehtovirta-Morley LE, Stoecker K, Vilcinskas A, Prosser JI, Nicol GW (2011). Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proc Natl Acad Sci USA* **108**: 15892-7.
- Leininger S, Urich T, Schlöter M, Schwark L, Qi J, Nicol GW *et al* (2006). *Archaea* predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**: 806-9.
- Lepère C (2007). *Diversité, dynamique et facteurs de régulation des picoeucaryotes dans les écosystèmes lacustres.*

- Lliros M, Casamayor EO, Borrego C (2008). High archaeal richness in the water column of a freshwater sulfurous karstic lake along an interannual study. *FEMS Microbiol Ecol* **66**: 331-342.
- Lliros M, Gich F, Plasencia A, Auguet JC, Darchambeau F, Casamayor EO *et al* (2010). Vertical distribution of ammonia-oxidizing *crenarchaeota* and methanogens in the epipelagic waters of Lake Kivu (Rwanda-Democratic Republic of the Congo). *Appl Environ Microbiol* **76**: 6853-63.
- Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD *et al* (2013). Predominant *archaea* in marine sediments degrade detrital proteins. *Nature* **496**: 215-8.
- Logue JB, Bürgmann H, Robinson CT (2008). Ecological Genetics and the Biodiversity of Freshwater Aquatic *Bacteria*. *BioScience* **52**: 103-113.
- Lorenzen CJ (1967a). Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnology and Oceanography* **12**: 343-346.
- Lorenzen CJ (1967b). Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnol. Oceanogr.* **12**: 343-346.
- Lozupone CA, Knight R (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104**: 11436-40.
- Lu L, Han W, Zhang J, Wu Y, Wang B, Lin X *et al* (2012). Nitrification of archaeal ammonia oxidizers in acid soils is supported by hydrolysis of urea. *ISME J* **6**: 1978-84.
- Lynch MD, Bartram AK, Neufeld JD (2012). Targeted recovery of novel phylogenetic diversity from next-generation sequence data. *ISME J* **6**: 2067-77.
- Mangot JF, Domaizon I, Taib N, Marouni N, Duffaud E, Bronner G *et al* (2012). Short-term dynamics of diversity patterns: evidence of continual reassembly within lacustrine small eukaryotes. *Environ Microbiol.*
- Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA (2009). Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**: 976-9.
- Martens-Habbena W, Stahl DA (2011). Nitrogen metabolism and kinetics of ammonia-oxidizing *archaea*. *Methods Enzymol* **496**: 465-87.
- Martin-Cuadrado AB, Rodriguez-Valera F, Moreira D, Alba JC, Ivars-Martinez E, Henn MR *et al* (2008). Hindsight in the relative abundance, metabolic potential and genome dynamics of uncultivated marine *archaea* from comparative metagenomic analyses of bathypelagic plankton of different oceanic regions. *ISME J* **2**: 865-86.
- Martiny JB, Bohannan BJ, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102-12.

- Mary I, Cummings DG, Biegala IC, Burkill PH, Archer SD, Zubkov MV (2006). Seasonal dynamics of bacterioplankton community structure at a coastal station in the western English Channel *Aqua Microbial Ecol* **42**: 119-126.
- Massana R, DeLong EF, Pedros-Alio C (2000). A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl Environ Microbiol* **66**: 1777-87.
- Matsumi R, Atomi H, Driessen AJ, van der Oost J (2010). Isoprenoid biosynthesis in *Archaea*--biochemical and evolutionary implications. *Res Microbiol* **162**: 39-52.
- Mayr E (1998). Two empires or three? *Proc Natl Acad Sci USA* **95**: 9720-3.
- Menendez C, Bauer Z, Huber H, Gad'on N, Stetter KO, Fuchs G (1999). Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic *Crenarchaeota* and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation. *J Bacteriol* **181**: 1088-98.
- Merbt SN, Stahl DA, Casamayor EO, Marti E, Nicol GW, Prosser JI (2012). Differential photoinhibition of bacterial and archaeal ammonia oxidation. *FEMS Microbiol Lett* **327**: 41-6.
- Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF (2007). Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ Microbiol* **9**: 1162-75.
- Moin NS, Nelson KA, Bush A, Bernhard AE (2009). Distribution and diversity of archaeal and bacterial ammonia oxidizers in salt marsh sediments. *Appl Environ Microbiol* **75**: 7461-8.
- Molina V, Belmar L, Ulloa O (2010). High diversity of ammonia-oxidizing *archaea* in permanent and seasonal oxygen-deficient waters of the eastern South Pacific. *Environ Microbiol* **12**: 2450-65.
- Moore LR, Rocap G, Chisholm SW (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464-7.
- Mosier AC, Allen EE, Kim M, Ferriera S, Francis CA (2012). Genome sequence of "Candidatus *Nitrosopumilus salaria*" BD31, an ammonia-oxidizing archaeon from the San Francisco Bay estuary. *J Bacteriol* **194**: 2121-2.
- Mosier AC, Francis CA (2008). Relative abundance and diversity of ammonia-oxidizing *archaea* and bacteria in the San Francisco Bay estuary. *Environ Microbiol* **10**: 3002-16.
- Mussmann M, Brito I, Pitcher A, Sinnighe Damste JS, Hatzenpichler R, Richter A *et al* (2011). Thaumarchaeotes abundant in refinery nitrifying sludges express *amoA* but are not obligate autotrophic ammonia oxidizers. *Proc Natl Acad Sci USA* **108**: 16771-16776.
- Nicol GW, Leininger S, Schleper C, Prosser JI (2008). The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing *archaea* and *bacteria*. *Environ Microbiol* **10**: 2966-78.

- Nicol GW, Schleper C (2006). Ammonia-oxidising *Crenarchaeota*: important players in the nitrogen cycle? *Trends Microbiol* **14**: 207-12.
- Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H *et al* (2011). Insights into the evolution of *Archaea* and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* **39**: 3204-23.
- Ochsenreiter T, Selezi D, Quaiser A, Bonch-Osmolovskaya L, Schleper C (2003). Diversity and abundance of *Crenarchaeota* in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environ Microbiol* **5**: 787-97.
- Offre P, Nicol GW, Prosser JI (2010). Community profiling and quantification of putative autotrophic thaumarchaeal communities in environmental samples. *Environmental Microbiology Reports*.
- Ogier S (1999). *Diagenèse précoce en domaine lacustre : étude des composés minéraux et organiques des sédiments récents du lac d'Aydat (Puy de Dôme, France)*.
- Ouverney CC, Fuhrman JA (2000). Marine planktonic *archaea* take up amino acids. *Appl Environ Microbiol* **66**: 4829-33.
- Pace NR (2006). Time for a change. *Nature* **441**: 289.
- Park BJ, Park SJ, Yoon DN, Schouten S, Sinninghe Damste JS, Rhee SK (2010). Cultivation of autotrophic ammonia-oxidizing *archaea* from marine sediments in coculture with sulfur-oxidizing *bacteria*. *Appl Environ Microbiol* **76**: 7575-87.
- Paynter MJ, Hungate RE (1968). Characterization of *Methanobacterium mobilis*, sp. n., isolated from the bovine rumen. *J Bacteriol* **95**: 1943-51.
- Pedros-Alio C (2006). Marine microbial diversity: can it be determined? *Trends Microbiol* **14**: 257-63.
- Pedros-Alio C (2012). The Rare Bacterial Biosphere. *Annu Rev Mar Sci* **4**: 449-466.
- Pernthaler J (2005). Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3**: 537-46.
- Pernthaler J, Glockner FO, Unterholzner S, Alfreider A, Psenner R, Amann R (1998). Seasonal community and population dynamics of pelagic *bacteria* and *archaea* in a high mountain lake. *Appl Environ Microbiol* **64**: 4299-306.
- Pester M, Schleper C, Wagner M (2011). The *Thaumarchaeota*: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol* **14**: 300-6.
- Pitcher A, Villanueva L, Hopmans EC, Schouten S, Reichart GJ, Sinninghe Damste JS (2011). Niche segregation of ammonia-oxidizing archaea and anammox bacteria in the Arabian Sea oxygen minimum zone. *ISME J*.

- Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA *et al* (2008). A genomic analysis of the archaeal system *Ignicoccus hospitalis*-*Nanoarchaeum equitans*. *Genome Biol* **9**: R158.
- Pouliot J, Galand PE, Lovejoy C, Vincent WF (2009). Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two meromictic High Arctic lakes. *Environ Microbiol* **11**: 687-99.
- Preston CM, Wu KY, Molinski TF, DeLong EF (1996). A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc Natl Acad Sci USA* **93**: 6241-6.
- Price MN, Dehal PS, Arkin AP (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.
- Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J *et al* (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188-96.
- Pulliam HR (1988). Sources, sinks, and population regulation. *The American Naturalist* **132**: 652-661.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590-6.
- Quince C, Lanzen A, Curtis TP, Davenport RJ, Hall N, Head IM *et al* (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Methods* **6**: 639-41.
- Ragan-Kelley B, Walters WA, McDonald D, Riley J, Granger BE, Gonzalez A *et al* (2012). Collaborative cloud-enabled tools allow rapid, reproducible biological insights. *ISME J*.
- Ramade F (2009). *Ecologie fondamentale*.
- Rysgaard S, Nielsen T, Hansen B (1999). Seasonal variation in nutrients, pelagic primary production and grazing in a high-Arctic coastal marine ecosystem, Young Sound, Northeast Greenland. *Marine ecology progress series* **179**: 13-25.
- Santoro AE, Casciotti K, Francis CA (2010). Activity, abundance and diversity of nitrifying archaea and bacteria in the central California Current. *Environmental Microbiology Reports* **12**: 1989-2006.
- Santoro AE, Francis CA, de Sieyes NR, Boehm AB (2008). Shifts in the relative abundance of ammonia-oxidizing *bacteria* and *archaea* across physicochemical gradients in a subterranean estuary. *Environ Microbiol* **10**: 1068-79.
- Schleper C, Nicol GW (2010). Ammonia-oxidising archaea--physiology, ecology and evolution. *Adv Microb Physiol* **57**: 1-41.

- Schneider D, Arp G, Reimer A, Reitner J, Daniel R (2013). Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the kiritimati atoll, central pacific. *PLoS One* **8**: e66662.
- Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ (2010). The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* **12**: 490-500.
- Simon J, Klotz MG (2012). Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim Biophys Acta*.
- Singh SK, Verma P, Ramaiah N, Chandrashekar AA, Shouche YS (2010). Phylogenetic diversity of archaeal 16S rRNA and ammonia monooxygenase genes from tropical estuarine sediments on the central west coast of India. *Res Microbiol* **161**: 177-86.
- Sintes E, Bergauer K, De Corte D, Yokokawa T, Herndl GJ (2013). Archaeal *amoA* gene diversity points to distinct biogeography of ammonia-oxidizing *Crenarchaeota* in the ocean. *Environ Microbiol* **15**: 1647-58.
- Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR *et al* (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* **103**: 12115-20.
- Spang A, Hatzenpichler R, Brochier-Armanet C, Rattei T, Tischler P, Spieck E *et al* (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum *Thaumarchaeota*. *Trends Microbiol* **18**: 331-40.
- Spang A, Poehlein A, Offre P, Zumbragel S, Haider S, Rychlik N *et al* (2012). The genome of the ammonia-oxidizing Candidatus *Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**: 3122-45.
- Stahl DA, de la Torre JR (2012). Physiology and diversity of ammonia-oxidizing *archaea*. *Annu Rev Microbiol* **66**: 83-101.
- Stewart FJ, Ulloa O, DeLong EF (2012). Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ Microbiol* **14**: 23-40.
- Strickland J, Parsons T (1968). *A practical handbook of sea water analysis*.
- Suzuki MT, Preston CM, Beja O, de la Torre JR, Steward GF, DeLong EF (2004). Phylogenetic screening of ribosomal RNA gene-containing clones in Bacterial Artificial Chromosome (BAC) libraries from different depths in Monterey Bay. *Microb Ecol* **48**: 473-88.
- Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ *et al* (2013). Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J* **7**: 509-519.

- Taib N, Mangot JF, Domaizon I, Bronner G, Debroas D (2013). Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS One* **8**: e58950.
- Takai K, Horikoshi K (2000). Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Appl Environ Microbiol* **66**: 5066-72.
- Takai K, Komatsu T, Inagaki F, Horikoshi K (2001a). Distribution of *Archaea* in a black smoker chimney structure. *Appl Environ Microbiol* **67**: 3618-29.
- Takai K, Moser DP, DeFlaun M, Onstott TC, Fredrickson JK (2001b). Archaeal diversity in waters from deep South African gold mines. *Appl Environ Microbiol* **67**: 5750-60.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731-9.
- Teira E, Reinthaler T, Pernthaler A, Pernthaler J, Herndl GJ (2004). Combining catalyzed reporter deposition-fluorescence in situ hybridization and microautoradiography to detect substrate utilization by bacteria and Archaea in the deep ocean. *Appl Environ Microbiol* **70**: 4411-4.
- Teske A (2006). Microbial communities of deep marine subsurface sediments: molecular and cultivation surveys. *Geomicrobiology Journal* **23**: 357-368.
- Teske A, Sorensen KB (2008). Uncultured *archaea* in deep marine subsurface sediments: have we caught them all? *ISME J* **2**: 3-18.
- Thauer RK (1998). Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144** (Pt 9): 2377-406.
- Tourna M, Freitag TE, Nicol GW, Prosser JI (2008). Growth, activity and temperature responses of ammonia-oxidizing *archaea* and *bacteria* in soil microcosms. *Environ Microbiol* **10**: 1357-64.
- Tourna M, Stieglmeier M, Spang A, Konneke M, Schintlmeister A, Urich T *et al* (2011). *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA* **108**: 8420-5.
- Treusch AH, Kletzin A, Raddatz G, Ochsenreiter T, Quaiser A, Meurer G *et al* (2004). Characterization of large-insert DNA libraries from soil for environmental genomic studies of Archaea. *Environ Microbiol* **6**: 970-80.
- Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic *crenarchaeota* in nitrogen cycling. *Environ Microbiol* **7**: 1985-95.
- Tully BJ, Nelson WC, Heidelberg JF (2012). Metagenomic analysis of a complex marine planktonic thaumarchaeal community from the Gulf of Maine. *Environ Microbiol* **14**: 254-67.

- Urakawa H, Martens-Habben W, Stahl D (2011). Physiology and Genomics of Ammonia-Oxidizing Archaea. In: Microbiology ASf (ed). *Nitrification*.
- Varela MM, van Aken HM, Sintes E, Reinthaler T, Herndl GJ (2011). Contribution of *Crenarchaeota* and *Bacteria* to autotrophy in the North Atlantic interior. *Environ Microbiol* **13**: 1524-33.
- Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al* (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.
- Viollier E, Michard G, Jezequel D, Pepe M, Sarazin G (1997). Geochemical study of a crater lake: Lake Pavin, Puy de Dôme, France. Constraints afforded by the particular matter distribution in the element cycling within the lake. *Chemical Geology* **142**: 225-241.
- Vissers EW, Anselmetti FS, Bodelier PL, Muyzer G, Schleper C, Tourna M *et al* (2013). Temporal and spatial coexistence of archaeal and bacterial amoA genes and gene transcripts in Lake Lucerne. *Archaea* **2013**: 289478.
- Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ *et al* (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107**: 8818-23.
- Wang P, Li T, Hu A, Wei Y, Guo W, Jiao N *et al* (2010). Community structure of archaea from deep-sea sediments of the South China Sea. *Microb Ecol* **60**: 796-806.
- Ward DM (2006). A macrobiological perspective on microbial species. *Microbe* **1**: 269-278.
- Whitaker RJ, Grogan DW, Taylor JW (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**: 976-8.
- Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV *et al* (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* **13**: 1310-24.
- Winter C, Bouvier T, Weinbauer MG, Thingstad TF (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiol Mol Biol Rev* **74**: 42-57.
- Winter C, Kerros ME, Weinbauer MG (2009). Seasonal changes of bacterial and archaeal communities in the dark ocean: Evidence from the Mediterranean Sea. *Limnol Oceanogr* **54**: 160-170.
- Woese CR (1994). There must be a prokaryote somewhere: microbiology's search for itself. *Microbiol Rev* **58**: 1-9.
- Woese CR, Fox GE (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA* **74**: 5088-90.

Woese CR, Kandler O, Wheelis ML (1990). Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* **87**: 4576-9.

Wuchter C, Abbas B, Coolen MJ, Herfort L, van Bleijswijk J, Timmers P *et al* (2006). Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* **103**: 12317-22.

Yakimov MM, Cono VL, Smedile F, Deluca TH, Juarez S, Ciordia S *et al* (2011). Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary production in Tyrrhenian deep waters (Central Mediterranean Sea). *ISME J* **5**: 945-61.

Yannarell AC, Triplett EW (2004). Within- and between-lake variability in the composition of bacterioplankton communities: investigations using multiple spatial scales. *Appl Environ Microbiol* **70**: 214-23.

CURRICULUM VITAE

CURRICULUM VITAE

Activités d'enseignement et de recherche

PUBLICATIONS

Revues internationales à comité de lecture

Publications acceptées

1. **2013 - Hugoni M.** & Taib N., Debroas D., Domaizon I., Jouan-Dufournel I., Bronner G., Salter I., Agogué H., Mary I., Galand PE. Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Science USA*. Doi : 10.1073/pnas.1216863110.
2. **2013 - Hugoni M.**, Etien S., Bourges A., Lepère C., Domaizon I., Mallet C., Bronner G., Debroas D., Mary I. Dynamics of ammonia-oxidizing Archaea and Bacteria in contrasted freshwater ecosystems. *Research in Microbiology*. Doi : 10.1016/j.resmic.2013.01.004.

Publications en préparation

1. **2013 – Hugoni M.**, Agogué H, Taib N, Domaizon I, Moné A, Galand P, Bronner G, Debroas D, Mary I. Temporal and spatial dynamics of active prokaryotes nitrifiers and archaeal communities from river to sea. *In prep.*
2. **2013 – Hugoni M.**, Domaizon I, Agogué H, Taib N, Moné A, Galand P, Bronner G, Debroas D, Mary I. New insights into active archaeal groups in oxygen depleted zones of lacustrine ecosystems. *In prep.*
3. 2013 – Taib N, Bronner G, Charvy JC, Roux S, **Hugoni M**, Mahul A, Mephu –Nguifo E, Breton V, Debroas D. ePANAM: a web server for depicting the microbial diversity from high-throughput sequencing amplicons. *In prep.*

CONGRES NATIONAUX ET INTERNATIONAUX

Oral

1. **2013 – Hugoni M.**, Debroas D., Agogu  H., Taib N., Domaizon I., Bronner G., Galand P.E., Mary I. Diversity and ecological role of mesophilic Archaea in aquatic ecosystems. *Journ es th matiques Archaea (Brest, France)*
2. **2013 - Hugoni M. & Taib N.**, Debroas D., Domaizon I., Jouan-Dufournel I., Bronner G., Salter I., Agogu  H., Mary I., Galand P.E. Study of archaeal dynamics in the Mediterranean Sea : new insights in ecological role of the rare biosphere. *Journ es de l'Ecole Doctorale Sciences de la Vie, de l'Environnement et de la Sant  (Clermont Ferrand, France)*
3. **2012 - Hugoni M. & Taib N.**, Debroas D., Domaizon I., Jouan-Dufournel I., Bronner G., Salter I., Agogu  H., Mary I., Galand P.E. Archaeal rare biosphere and seasonal dynamics of active ecotypes in coastal waters. *Journ es Internationales de Limnologie et d'Oc anographie (JILO) (Clermont Ferrand, France)*
4. **2012 - Roux S.**, Taib N., Mangot J.F., **Hugoni M.**, Mary I., Ravet V., Bronner G., Enault F., Debroas D. High-throughput sequencing data analysis through phylogenetic approaches: tools and methods. *Colloque de G nomique Environnementale (Lyon, France)*
5. **2011 - Hugoni M.**, Etien S., Bourges A., Lep re C., Domaizon I., Mallet C., Bronner G., Debroas D., Mary I. Diversity and spatio-temporal dynamics of Archaea and ammonia oxidizing communities in the photic zone of temperate lakes. *12 me Symposium of Aquatic and Microbial Ecology (SAME 12), Rostock, Germany*
6. **2011 - Hugoni M.**, Etien S., Bourges A., Lep re C., Domaizon I., Mallet C., Bronner G., Debroas D., Mary I. Structure and temporal dynamics of ammonia oxidizing Archaea in lacustrine ecosystems. *8 mes Rencontres des Microbiologistes Clermontois (Clermont Ferrand, France)*

Poster

1. **2013 – Hugoni M.**, Agogu  H., Taib N., Domaizon I., Mon  A., Galand P., Bronner G., Debroas D., Mary I. Temporal and spatial dynamics of active prokaryotes nitrifiers and archaeal communities from river to sea. *13 me Symposium of Aquatic and Microbial Ecology (SAME 13), Stresa, Italy*
2. **2013 – Hugoni M.**, Agogu  H., Taib N., Domaizon I., Mon  A., Galand P., Bronner G., Debroas D., Mary I. Temporal changes in archaeal community composition associated with Ammonia Oxidizing Archaea dynamics along an estuarine salinity gradient. *Journ es th matiques Archaea (Brest, France)*
3. **2011 - Hugoni M.**, Etien S., Bourges A., Lep re C., Domaizon I., Mallet C., Bronner G., Debroas D., Mary I. Community structure and temporal dynamics of ammonia-

oxidizing Crenarchaeota in the photic zone of French lakes. E Colloque de Génomique Environnementale (Lyon, France)

4. **2011 - Hugoni M.**, Etien S., Bourges A., Lepère C., Domaizon I., Mallet C., Bronner G., Debroas D., Mary I. Community structure and temporal dynamics of ammonia-oxidizing Crenarchaeota in the photic zone of French lakes. Journées de l'Ecole Doctorale Sciences de la Vie, de l'Environnement et de la Santé (Clermont Ferrand, France, Poster primé)

PARTICIPATION A DES PROJETS DE RECHERCHE

Les recherches menées au cours de mon doctorat se sont inscrites dans le cadre de plusieurs projets :

Programme EC2CO DIVAQUA. *Diversity and ecological importance of mesophilic Archaea in ammonia oxidation in aquatic ecosystems* (2010-2012). PI : I. Mary.

Projet PICS CNRS: *Diversity and biogeochemical role of Archaea in aquatic ecosystems* (2010-2013). PI: I. Mary.

Programme APEGE: *MetaArch.* (2012-2013). PI: I. Mary.

SOERE GLACPE: Observatoire des grands lacs périalpins

ACTIVITES d'ENSEIGNEMENT

2010-2013 : Monitrice de l'Université Blaise Pascal, Clermont Ferrand II

Enseignements dispensés : TP de Biochimie et Immunologie, étudiants de première et seconde année en Génie Biologique (192h)

Etablissement : Ecole d'Ingénieurs du réseau Polytech'

Responsable de l'Enseignement : Guillaume PIERRE et Marièle BRIAND

